

明 細 書

音声・楽音符号化装置及び音声・楽音符号化方法

技術分野

- [0001] 本発明は、インターネット通信に代表されるパケット通信システムや、移動通信システムなどで音声・楽音信号の伝送を行う音声・楽音符号化装置及び音声・楽音符号化方法に関する。

背景技術

- [0002] インターネット通信に代表されるパケット通信システムや、移動通信システムなどで音声信号を伝送する場合、伝送効率を高めるために圧縮・符号化技術が利用される。これまでに多くの音声符号化方式が開発され、近年開発された低ビットレート音声符号化方式の多くは、音声信号をスペクトル情報とスペクトルの微細構造情報とに分離し、分離したそれぞれに対して圧縮・符号化を行うという方式である。
- [0003] また、IP電話に代表されるようなインターネット上での音声通話環境が整備されつつあり、音声信号を効率的に圧縮して転送する技術に対するニーズが高まっている。
- [0004] 特に、人間の聴感マスキング特性を利用した音声符号化に関する様々な方式が検討されている。聴感マスキングとは、ある周波数に含まれる強い信号成分が存在する時に、隣接する周波数成分が、聞こえなくなる現象でこの特性を利用して品質向上を図るものである。
- [0005] これに関連した技術としては、例えば、ベクトル量子化の距離計算時に聴感マスキング特性を利用した特許文献1に記載されるような方法がある。
- [0006] 特許文献1の聴感マスキング特性を用いた音声符号化手法は、入力された信号の周波数成分と、コードブックが示すコードベクトルの双方が聴感マスキング領域にある場合、ベクトル量子化時の距離を0とする計算方法である。これにより、聴感マスキング領域外における距離の重みが相対的に大きくなり、より効率的に音声符号化することが可能となる。

特許文献1:特開平8-123490号公報(第3頁、第1図)

発明の開示

発明が解決しようとする課題

- [0007] しかしながら、特許文献1に示す従来方法では、入力信号及びコードベクトルの限られた場合にしか適応できず音質性能が不十分であった。
- [0008] 本発明の目的は、上記の課題に鑑みてなされたものであり、聴感的に影響の大きい信号の劣化を抑える適切なコードベクトルを選択し、高品質な音声・楽音符号化装置及び音声・楽音符号化方法を提供することである。

課題を解決するための手段

- [0009] 上記課題を解決するために、本発明の音声・楽音符号化装置は、音声・楽音信号を時間成分から周波数成分へ変換する直交変換処理手段と、前記音声・楽音信号から聴感マスキング特性値を求める聴感マスキング特性値算出手段と、前記聴感マスキング特性値に基づいて、前記周波数成分と、予め設定されたコードブックから求めたコードベクトルと前記周波数成分と間の距離計算方法を変えてベクトル量子化を行うベクトル量子化手段と、を具備する構成を採る。

発明の効果

- [0010] 本発明によれば、聴感マスキング特性値に基づき、入力信号とコードベクトルとの距離計算方法を変えて量子化を行うことにより、聴感的に影響の大きい信号の劣化を抑える適切なコードベクトルを選択することが可能になり、入力信号の再現性を高め良好な復号化音声を得ることができる。

図面の簡単な説明

- [0011] [図1]本発明の実施の形態1に係る音声・楽音符号化装置及び音声・楽音復号化装置を含むシステム全体のブロック構成図
- [図2]本発明の実施の形態1に係る音声・楽音符号化装置のブロック構成図
- [図3]本発明の実施の形態1に係る聴感マスキング特性値算出部のブロック構成図
- [図4]本発明の実施の形態1に係る臨界帯域幅の構成例を示す図
- [図5]本発明の実施の形態1に係るベクトル量子化部のフローチャート
- [図6]本発明の実施の形態1に係る聴感マスキング特性値と符号化値とMDCT係数の相対的位置関係を説明する図

[図7]本発明の実施の形態1に係る音声・楽音復号化装置のブロック構成図

[図8]本発明の実施の形態2に係る音声・楽音符号化装置及び音声・楽音復号化装置のブロック構成図

[図9]本発明の実施の形態2に係るCELP方式の音声符号化装置の構成概要図

[図10]本発明の実施の形態2に係るCELP方式の音声復号化装置の構成概要図

[図11]本発明の実施の形態2に係る拡張レイヤ符号化部のブロック構成図

[図12]本発明の実施の形態2に係るベクトル量子化部のフローチャート

[図13]本発明の実施の形態2に係る聴感マスキング特性値と符号化値とMDCT係数の相対的位置関係を説明する図

[図14]本発明の実施の形態2に係る復号化部のブロック構成図

[図15]本発明の実施の形態3に係る音声信号送信装置及び音声信号受信装置のブロック構成図

[図16]本発明の実施の形態1に係る符号化部のフローチャート

[図17]本発明の実施の形態1に係る聴感マスキング値算出部のフローチャート

発明を実施するための最良の形態

[0012] 以下、本発明の実施の形態について、添付図面を参照して詳細に説明する。

[0013] (実施の形態1)

図1は、本発明の実施の形態1に係る音声・楽音符号化装置及び音声・楽音復号化装置を含むシステム全体の構成を示すブロック図である。

[0014] このシステムは、入力信号を符号化する音声・楽音符号化装置101と伝送路103と受信した信号を復号化する音声・楽音復号化装置105から構成される。

[0015] なお、伝送路103は、無線LANあるいは携帯端末のパケット通信、Bluetoothなどの無線伝送路であってもよいし、ADSL、FTTHなどの有線伝送路であってもよい。

[0016] 音声・楽音符号化装置101は、入力信号100を符号化し、その結果を符号化情報102として伝送路103に出力する。

[0017] 音声・楽音復号化装置105は、伝送路103を介して符号化情報102を受信し、復号化し、その結果を出力信号106として出力する。

[0018] 次に、音声・楽音符号化装置101の構成について図2のブロック図を用いて説明す

る。図2において、音声・楽音符号化装置101は、入力信号100を時間成分から周波数成分へ変換する直交変換処理部201と、入力信号100から聴感マスキング特性値を算出する聴感マスキング特性値算出部203と、インデックスと正規化されたコードベクトルの対応を示す形状コードブック204と、形状コードブック204の正規化された各コードベクトルに対応してその利得を示す利得コードブック205と、前記聴感マスキング特性値、前記形状コードブック及び利得コードブックを用いて前記周波数成分へ変換された入力信号をベクトル量子化するベクトル量子化部202とから主に構成される。

[0019] 次に、図16のフローチャートの手順に従って、音声・楽音符号化装置101の動作について、詳細に説明する。

[0020] まず、入力信号のサンプリング処理について説明する。音声・楽音符号化装置101は、入力信号100をNサンプルずつ区切り(Nは自然数)、Nサンプルを1フレームとしてフレーム毎に符号化を行う。ここで、符号化の対象となる入力信号100を x_n ($n=0, \Lambda, N-1$)と表すこととする。 n は前記区切られた入力信号である信号要素の $n+1$ 番目であることを示す。

[0021] 入力信号 x_n 100は、直交変換処理部201及び聴感マスキング特性算出部203に入力される。

[0022] 次に、直交変換処理部201は、前記信号要素に対応してバッファ buf_n ($n=0, \Lambda, N-1$)を内部に有し、式(1)によりそれぞれ0を初期値として初期化する。

[0023] [数1]

$$buf_n = 0 \quad (n=0, \dots, N-1) \quad \dots (1)$$

[0024] 次に、直交変換処理(ステップS1601)について、直交変換処理部201における計算手順と内部バッファへのデータ出力に関して説明する。

[0025] 直交変換処理部201は、入力信号 x_n 100を修正離散コサイン変換(MDCT)し、式(2)によりMDCT係数 X_k を求める。

[0026] [数2]

$$X_k = \frac{2}{N} \sum_{n=0}^{2N-1} x'_n \cos \left[\frac{(2n+1+N)(2k+1)\pi}{4N} \right] \quad (k=0, \dots, N-1) \quad \dots (2)$$

[0027] ここで、 k は1フレームにおける各サンプルのインデックスを意味する。直交変換処理部201は、入力信号 x_n 100とバッファ buf_n とを結合させたベクトルである x'_n を式(3)により求める。

[0028] [数3]

$$x'_n = \begin{cases} buf_n & (n=0, \dots, N-1) \\ x_{n-N} & (n=N, \dots, 2N-1) \end{cases} \quad \dots (3)$$

[0029] 次に、直交変換処理部201は、式(4)によりバッファ buf_n を更新する。

[0030] [数4]

$$buf_n = x_n \quad (n=0, \dots, N-1) \quad \dots (4)$$

[0031] 次に、直交変換処理部201は、MDCT係数 X_k をベクトル量子化部202に出力する。

[0032] 次に、図2の聴感マスキング特性値算出部203の構成について、図3のブロック図を用いて説明する。

[0033] 図3において、聴感マスキング特性値算出部203は、入力信号をフーリエ変換するフーリエ変換部301と、前記フーリエ変換された入力信号からパワースペクトルを算出するパワースペクトル算出部302と、入力信号から最小可聴閾値を算出する最小可聴閾値算出部304と、前記算出された最小可聴閾値をバッファリングするメモリバッファ305と、前記算出されたパワースペクトルと前記バッファリングされた最小可聴閾値から聴感マスキング値を計算する聴感マスキング値算出部303とから構成される。

[0034] 次に、上記のように構成された聴感マスキング特性値算出部203における聴感マスキング特性値算出処理(ステップS1602)について、図17のフローチャートを用いて動作を説明する。

[0035] なお、聴感マスキング特性値の算出方法については、Johnston氏らによる論文(J. Johnston, "Estimation of perceptual entropy using noise masking criteria", in Proc. ICASSP-88, May 1988, pp.2524-2527)に開示されている。

[0036] まず、フーリエ変換処理(ステップS1701)についてフーリエ変換部301の動作を説明する。

[0037] フーリエ変換部301は、入力信号 x_n 100を入力して、これを式(5)により周波数領域の信号 F_k に変換する。ここで、 e は自然対数の底であり、 k は1フレームにおける各サンプルのインデックスである。

[0038] [数5]

$$F_k = \sum_{n=0}^{N-1} x_n e^{-j \frac{2\pi kn}{N}} \quad (k=0, \dots, N-1) \quad \dots (5)$$

[0039] 次に、フーリエ変換部301は、得られた F_k をパワースペクトル算出部302に出力する。

[0040] 次に、パワースペクトル算出処理(ステップS1702)について説明する。

[0041] パワースペクトル算出部302は、前記フーリエ変換部301から出力された周波数領域の信号 F_k を入力とし、式(6)により F_k のパワースペクトル P_k を求める。ただし、 k は1フレームにおける各サンプルのインデックスである。

[0042] [数6]

$$P_k = (F_k^{\text{Re}})^2 + (F_k^{\text{Im}})^2 \quad (k=0, \dots, N-1) \quad \dots (6)$$

[0043] なお、式(6)において、 F_k^{Re} は周波数領域の信号 F_k の実部であり、パワースペクトル算出部302は、式(7)により F_k^{Re} を求める。

[0044] [数7]

$$F_k^{\text{Re}} = \sum_{n=0}^{N-1} \left[x_n \cos\left(\frac{2\pi kn}{N}\right) \right] \quad (k=0, \dots, N-1) \quad \dots (7)$$

[0045] また、 F_k^{Im} は周波数領域の信号 F_k の虚部であり、パワースペクトル算出部302は、式(8)により F_k^{Im} を求める。

[0046] [数8]

$$F_k^{\text{Im}} = -\sum_{n=0}^{N-1} \left[x_n \sin\left(\frac{2\pi kn}{N}\right) \right] \quad (k=0, \dots, N-1) \quad \dots (8)$$

[0047] 次に、パワースペクトル算出部302は、得られたパワースペクトル P_k を聴感マスキング値算出部303に出力する。

[0048] 次に、最小可聴閾値算出処理(ステップS1703)について説明する。

[0049] 最小可聴閾値算出部304は、第1フレームにおいてのみ、式(9)により最小可聴閾

値 ath_k を求める。

[0050] [数9]

$$ath_k = 3.64(k/1000)^{-0.8} - 6.5e^{-0.6(k/1000-3.3)^2} + 10^{-3}(k/1000)^4 \quad (k=0, \dots, N-1) \quad \dots (9)$$

[0051] 次に、メモリバッファへの保存処理(ステップS1704)について説明する。

[0052] 最小可聴閾値算出部304は、最小可聴閾値 ath_k をメモリバッファ305に出力する。メモリバッファ305は、入力された最小可聴閾値 ath_k を聴感マスキング値算出部303に出力する。最小可聴閾値 ath_k とは、人間の聴覚に基づき各周波数成分に対して定められ、 ath_k 以下の成分は聴感的に知覚することができないという値である。

[0053] 次に、聴感マスキング値算出処理(ステップS1705)について聴感マスキング値算出部303の動作を説明する。

[0054] 聴感マスキング値算出部303は、パワースペクトル算出部302から出力されたパワースペクトル P_k を入力し、パワースペクトル P_k を m の臨界帯域幅に分割する。ここで、臨界帯域幅とは、帯域雑音を増加してもその中心周波数の純音がマスクされる量が増えなくなる限界の帯域幅のことである。また、図4に、臨界帯域幅の構成例を示す。図4において、 m は臨界帯域幅の総数であり、パワースペクトル P_k は m の臨界帯域幅に分割される。また、 i は臨界帯域幅のインデックスであり、 $0 \sim m-1$ の値をとる。また、 bh_i 及び bl_i は各臨界帯域幅 i の最小周波数インデックス及び最大周波数インデックスである。

[0055] 次に、聴感マスキング値算出部303は、パワースペクトル算出部302から出力されたパワースペクトル P_k を入力し、式(10)により臨界帯域幅毎に加算されたパワースペクトル B_i を求める。

[0056] [数10]

$$B_i = \sum_{k=bl_i}^{bh_i} P_k \quad (i=0, \dots, m-1) \quad \dots (10)$$

[0057] 次に、聴感マスキング値算出部303は、式(11)により拡散関数 $SF(t)$ (Spreading Function)を求める。拡散関数 $SF(t)$ とは、各周波数成分に対して、その周波数成分が近隣周波数に及ぼす影響(同時マスキング効果)を算出するために用いるものである。

[0058] [数11]

$$SF(t) = 15.81139 + 7.5(t + 0.474) - 17.5\sqrt{1 + (t + 0.474)^2} \quad (t = 0, \dots, N_i - 1) \quad \dots (11)$$

[0059] ここで、 N_i は定数であり、式(12)の条件を満たす範囲内で予め設定される。

[0060] [数12]

$$0 \leq N_i \leq m \quad \dots (12)$$

[0061] 次に、聴感マスキング値算出部303は、式(13)により臨界帯域幅毎に加算されたパワースペクトル B_i と拡散関数 $SF(t)$ を用い、定数 C_i を求める。

[0062] [数13]

$$C_i = \begin{cases} \sum_{t=N_T-i}^{N_i} B_i \cdot SF(t) & (i < N_i) \\ \sum_{t=0}^{N_i} B_i \cdot SF(t) & (N_i \leq i \leq N - N_i) \\ \sum_{t=0}^{N-i} B_i \cdot SF(t) & (i > N - N_i) \end{cases} \quad \dots (13)$$

[0063] 次に、聴感マスキング値算出部303は、式(14)により幾何平均 μ_i^g を求める。

[0064] [数14]

$$\mu_i^g = 10^{\frac{\log \left(\prod_{k=bh_i}^{bl_i} P_k \right)}{bl_i - bh_i}} \quad (i = 0, \dots, m-1) \quad \dots (14)$$

[0065] 次に、聴感マスキング値算出部303は、式(15)により算術平均 μ_i^a を求める。

[0066] [数15]

$$\mu_i^a = \sum_{k=bh_i}^{bl_i} P_k / (bl_i - bh_i) \quad (i = 0, \dots, m-1) \quad \dots (15)$$

[0067] 次に、聴感マスキング値算出部303は、式(16)により SFM_i (Spectral Flatness Measure)を求める。

[0068] [数16]

$$SFM_i = \mu_i^g / \mu_i^a \quad (i = 0, \dots, m-1) \quad \dots (16)$$

[0069] 次に、聴感マスキング値算出部303は、式(17)により定数 α_i を求める。

[0070] [数17]

$$\alpha_i = \min\left(\frac{10 \cdot \log_{10} SFM_i}{-60}, 1\right) \quad (i=0, \dots, m-1) \quad \dots (17)$$

[0071] 次に、聴感マスキング値算出部303は、式(18)により臨界帯域幅毎のオフセット値 O_i を求める。

[0072] [数18]

$$O_i = \alpha_i \cdot (14.5 + i) + 5.5 \cdot (1 - \alpha_i) \quad (i=0, \dots, m-1) \quad \dots (18)$$

[0073] 次に、聴感マスキング値算出部303は、式(19)により臨界帯域幅毎の聴感マスキング値 T_i を求める。

[0074] [数19]

$$T_i = \sqrt{10^{\log_{10}(C_i) - (O_i/10)} / (bl_i - bh_i)} \quad (i=0, \dots, m-1) \quad \dots (19)$$

[0075] 次に、聴感マスキング値算出部303は、メモリバッファ305から出力される最小可聴閾値 ath_k から、式(20)により聴感マスキング特性値 M_k を求め、これをベクトル量子化部202に出力する。

[0076] [数20]

$$M_k = \max(ath_k, T_i) \quad (k=bh_i, \dots, bl_i, i=0, \dots, m-1) \quad \dots (20)$$

[0077] 次に、ベクトル量子化部202における処理であるコードブック取得処理(ステップS1603)及びベクトル量子化処理(ステップS1604)について、図5処理フローを用いて詳細に説明する。

[0078] ベクトル量子化部202は、直交変換処理部201から出力されるMDCT係数 X_k と前記聴感マスキング特性値算出部203から出力される聴感マスキング特性値から、形状コードブック204、及び利得コードブック205を用いて、MDCT係数 X_k のベクトル量子化を行い、得られた符号化情報102を、図1の伝送路103に出力する。

[0079] 次に、コードブックについて説明する。

[0080] 形状コードブック204は、予め作成された N 種類の N 次元コードベクトル $code_j^l$ ($j=0, \Lambda, N_j-1, k=0, \Lambda, N-1$)から構成され、また、利得コードブック205は、予め作成された N_d 種類の利得コード $gain_d$ ($j=0, \Lambda, N_d-1$)から構成される。

[0081] ステップ501では、形状コードブック204におけるコードベクトルインデックス j に0を代入し、最小誤差 $Dist_{MIN}$ に十分大きな値を代入し、初期化する。

[0082] ステップ502では、形状コードブック204からN次元のコードベクトル $code_k^j$ ($k=0, \Lambda, N-1$)を読み込む。

[0083] ステップ503では、直交変換処理部201から出力されたMDCT係数 X_k を入力して、ステップ502の形状コードブック204で読み込んだコードベクトル $code_k^j$ ($k=0, \Lambda, N-1$)の利得Gainを式(21)により求める。

[0084] [数21]

$$Gain = \sum_{k=0}^{N-1} X_k \cdot code_k^j / \sum_{k=0}^{N-1} code_k^{j^2} \quad \dots (21)$$

[0085] ステップ504では、ステップ505の実行回数を表すcalc_countに0を代入する。

[0086] ステップ505では、聴感マスキング特性値算出部203から出力された聴感マスキング特性値 M_k を入力し、式(22)により一時利得 $temp_k$ ($k=0, \Lambda, N-1$)を求める。

[0087] [数22]

$$temp_k = \begin{cases} code_k^j & (|code_k^j \cdot Gain| \geq M_k) \\ 0 & (|code_k^j \cdot Gain| < M_k) \end{cases} \quad (k=0, \dots, N-1) \quad \dots (22)$$

[0088] なお、式(22)において、 k が $|code_k^j \cdot Gain| \geq M_k$ の条件を満たす場合、一時利得 $temp_k$ には $code_k^j$ が代入され、 k が $|code_k^j \cdot Gain| < M_k$ の条件を満たす場合、一時利得 $temp_k$ には0が代入される。

[0089] 次に、ステップ505では、式(23)により聴感マスキング値以上の要素に対する利得Gainを求める。

[0090] [数23]

$$Gain = \sum_{k=0}^{N-1} X_k \cdot temp_k / \sum_{k=0}^{N-1} temp_k^2 \quad (k=0, \dots, N-1) \quad \dots (23)$$

[0091] ここで、全ての k において一時利得 $temp_k$ が0の場合には利得Gainに0を代入する。また、式(24)により、利得Gainと $code_k^j$ から符号化値 R_k を求める。

[0092] [数24]

$$R_k = Gain \cdot code_k^j \quad (k=0, \dots, N-1) \quad \dots (24)$$

[0093] ステップ506では、calc_countに1を足し加える。

[0094] ステップ507では、calc_countと予め定められた非負の整数 N_c とを比較し、calc

__countが N_c より小さい値である場合はステップ505に戻り、calc__countが N_c 以上である場合はステップ508に進む。このように、利得Gainを繰り返し求めることにより、利得Gainを適切な値にまで収束させることができる。

[0095] ステップ508では、累積誤差Distに0を代入し、また、サンプルインデックスkに0を代入する。

[0096] 次に、ステップ509、511、512、及び514において、聴感マスキング特性値 M_k と符号化値 R_k とMDCT係数 X_k との相対的な位置関係について場合分けを行い、場合分けの結果に応じてそれぞれステップ510、513、515、及び516で距離計算を行う。

[0097] この相対的な位置関係による場合分けを図6に示す。図6において、白い丸記号(○)は入力信号のMDCT係数 X_k を意味し、黒い丸記号(●)は符号化値 R_k を意味する。また、図6に示したものが本発明の特徴を示しているもので、聴感マスキング特性値算出部203で求めた聴感マスキング特性値 $+M_k - 0 - M_k$ の領域を聴感マスキング領域と呼び、入力信号のMDCT係数 X_k または符号化値 R_k がこの聴感マスキング領域に存在する場合の距離計算の方法を変えて計算することにより、より聴感的に近い高品質な結果を得ることができる。

[0098] ここで、図6を用いて、本発明におけるベクトル量子化時の距離計算法について説明する。図6の「場合1」に示すように入力信号のMDCT係数 X_k (○)と符号化値 R_k (●)のいずれかも聴感マスキング領域に存在せず、かつMDCT係数 X_k と符号化値 R_k とが同符号である場合には入力信号のMDCT係数 X_k (○)と符号化値 R_k (●)の距離 D_{11} を単純に計算する。また、図6の「場合3」、「場合4」に示すように入力信号のMDCT係数 X_k (○)と符号化値 R_k (●)のいずれかが聴感マスキング領域に存在する場合には、聴感マスキング領域内の位置を M_k 値(場合によっては、 $-M_k$ 値)に補正して D_{31} または D_{41} として計算する。また、図6の「場合2」に示すように入力信号のMDCT係数 X_k (○)と符号化値 R_k (●)が聴感マスキング領域をまたがって存在する場合には、聴感マスキング領域間の距離を $\beta \cdot D_{23}$ (β は任意の係数)と計算する。図6の「場合5」に示すように入力信号のMDCT係数 X_k (○)と符号化値 R_k (●)が共に聴感マスキング領域内に存在する場合には、距離 $D_{51} = 0$ として計算する。

[0099] 次に、ステップ509～ステップ517の各場合における処理について説明する。

[0100] ステップ509では、聴感マスキング特性値 M_k と符号化値 R_k とMDCT係数 X_k との相対的な位置関係が図6における「場合1」に該当するかどうかを式(25)の条件式により判定する。

[0101] [数25]

$$(|X_k| \geq M_k) \text{ and } (|R_k| \geq M_k) \text{ and } (X_k \cdot R_k \geq 0) \quad \dots (25)$$

[0102] 式(25)は、MDCT係数 X_k の絶対値と符号化値 R_k の絶対値とが共に聴感マスキング特性値 M_k 以上であり、かつ、MDCT係数 X_k と符号化値 R_k とが同符号である場合を意味する。聴感マスキング特性値 M_k とMDCT係数 X_k と符号化値 R_k とが式(25)の条件式を満たした場合は、ステップ510に進み、式(25)の条件式を満たさない場合は、ステップ511に進む。

[0103] ステップ510では、式(26)により符号化値 R_k とMDCT係数 X_k との誤差 $Dist_1$ を求め、累積誤差 $Dist$ に誤差 $Dist_1$ を加算し、ステップ517に進む。

[0104] [数26]

$$\begin{aligned} Dist_1 &= D_{11} \\ &= |X_k - R_k| \end{aligned} \quad \dots (26)$$

[0105] ステップ511では、聴感マスキング特性値 M_k と符号化値 R_k とMDCT係数 X_k との相対的な位置関係が図6における「場合5」に該当するかどうかを式(27)の条件式により判定する。

[0106] [数27]

$$(|X_k| \geq M_k) \text{ and } (|R_k| \geq M_k) \text{ and } (X_k \cdot R_k < 0) \quad \dots (27)$$

[0107] 式(27)は、MDCT係数 X_k の絶対値と符号化値 R_k の絶対値とが共に聴感マスキング特性値 M_k 以下である場合を意味する。聴感マスキング特性値 M_k とMDCT係数 X_k と符号化値 R_k とが式(27)の条件式を満たした場合は、符号化値 R_k とMDCT係数 X_k との誤差は0とし、累積誤差 $Dist$ には何も加算せずにステップ517に進み、式(27)の条件式を満たさない場合は、ステップ512に進む。

[0108] ステップ512では、聴感マスキング特性値 M_k と符号化値 R_k とMDCT係数 X_k との相対的な位置関係が図6における「場合2」に該当するかどうかを式(28)の条件式により判定する。

[0109] [数28]

$$Dist_2 = D_{21} + D_{22} + \beta * D_{23} \quad \dots (28)$$

[0110] 式(28)は、MDCT係数 X_k の絶対値と符号化値 R_k の絶対値とが共に聴感マスキング特性値 M_k 以上であり、かつ、MDCT係数 X_k と符号化値 R_k とが異符号である場合を意味する。聴感マスキング特性値 M_k とMDCT係数 X_k と符号化値 R_k とが式(28)の条件式を満たした場合は、ステップ513に進み、式(28)の条件式を満たさない場合は、ステップ514に進む。

[0111] ステップ513では、式(29)により符号化値 R_k とMDCT係数 X_k との誤差 $Dist_2$ を求め、累積誤差 $Dist$ に誤差 $Dist_2$ を加算し、ステップ517に進む。

[0112] [数29]

$$D_{21} = |X_k| - M_k \quad \dots (29)$$

[0113] ここで、 β は、MDCT係数 X_k 、符号化値 R_k 及び聴感マスキング特性値 M_k に応じて適宜設定される値であり、1以下の値が適当であり、被験者の評価により実験的に求めた数値を採用してもよい。また、 D_{21} 、 D_{22} 及び D_{23} は、それぞれ式(30)、式(31)及び式(32)により求める。

[0114] [数30]

$$D_{22} = |R_k| - M_k \quad \dots (30)$$

[0115] [数31]

$$D_{23} = M_k \cdot 2 \quad \dots (31)$$

[0116] [数32]

$$(|X_k| \geq M_k) \text{ and } (|R_k| < M_k) \quad \dots (32)$$

[0117] ステップ514では、聴感マスキング特性値 M_k と符号化値 R_k とMDCT係数 X_k との相対的な位置関係が図6における「場合3」に該当するかどうかを式(33)の条件式により判定する。

[0118] [数33]

$$\begin{aligned} Dist_3 &= D_{31} \\ &= |X_k| - M_k \end{aligned} \quad \dots (33)$$

[0119] 式(33)は、MDCT係数 X_k の絶対値が聴感マスキング特性値 M_k 以上であり、かつ、符号化値 R_k が聴感マスキング特性値 M_k 未満である場合を意味する。聴感マスキング特性値 M_k とMDCT係数 X_k と符号化値 R_k とが式(33)の条件式を満たした場合は、ステップ515に進み、式(33)の条件式を満たさない場合は、ステップ516に進む。

[0120] ステップ515では、式(34)により符号化値 R_k とMDCT係数 X_k との誤差 $Dist_3$ を求め、累積誤差 $Dist$ に誤差 $Dist_3$ を加算し、ステップ517に進む。

[0121] [数34]

$$(|X_k| < M_k) \text{ and } (|R_k| \geq M_k) \quad \dots (34)$$

[0122] ステップ516は、聴感マスキング特性値 M_k と符号化値 R_k とMDCT係数 X_k との相対的な位置関係が図6における「場合4」に該当し、式(35)の条件式を満たす。

[0123] [数35]

$$(|X_k| < M_k) \text{ and } (|R_k| < M_k) \quad \dots (35)$$

[0124] 式(35)は、MDCT係数 X_k の絶対値が聴感マスキング特性値 M_k 未満であり、かつ、符号化値 R_k が聴感マスキング特性値 M_k 以上である場合を意味する。この時、ステップ516では、式(36)により符号化値 R_k とMDCT係数 X_k との誤差 $Dist_4$ を求め、累積誤差 $Dist$ に誤差 $Dist_4$ を加算し、ステップ517に進む。

[0125] [数36]

$$\begin{aligned} Dist_4 &= D_{41} \\ &= |R_k| - M_k \end{aligned} \quad \dots (36)$$

[0126] ステップ517では、 k に1を足し加える。

[0127] ステップ518では、 N と k を比較し、 k が N より小さい値の場合は、ステップ509に戻る。 k が N と同じ値の場合は、ステップ519に進む。

[0128] ステップ519では、累積誤差 $Dist$ と最小誤差 $Dist_{MIN}$ とを比較し、累積誤差 $Dist$ が最小誤差 $Dist_{MIN}$ より小さい値の場合は、ステップ520に進み、累積誤差 $Dist$ が最小誤差 $Dist_{MIN}$ 以上である場合は、ステップ521に進む。

[0129] ステップ520では、最小誤差 $Dist_{MIN}$ に累積誤差 $Dist$ を代入し、 $code_index_{MIN}$ に j を代入し、誤差最小利得 $Dist_{MIN}$ に利得 $Gain$ を代入し、ステップ521に進む。

[0130] ステップ521では、 j に1を足し加える。

[0131] ステップ522では、コードベクトルの総数 N_j と j とを比較し、 j が N_j より小さい値の場合は、ステップ502に戻る。 j が N_j 以上である場合は、ステップ523に進む。

[0132] ステップ523では、利得コードブック205から N_d 種類の利得コード $gain^d$ ($d=0, \Lambda, N_d-1$)を読み込み、全ての d に対して式(37)により量子化利得誤差 $gainerr^d$ ($d=0, \Lambda, N_d-1$)を求める。

[0133] [数37]

$$gainerr^d = |Gain_{MIN} - gain^d| \quad (d = 0, \dots, N_d - 1) \quad \dots (37)$$

[0134] 次に、ステップ523では、量子化利得誤差 $gainerr^d$ ($d=0, \Lambda, N_d-1$)を最小とする d を求め、求めた d を $gain_index_{MIN}$ に代入する。

[0135] ステップ524では、累積誤差Distが最小となるコードベクトルのインデックスである $code_index_{MIN}$ とステップ523で求めた $gain_index_{MIN}$ とを符号化情報102として、図1の伝送路103に出力し、処理を終了する。

[0136] 以上が、符号化部101の処理の説明である。

[0137] 次に、図1の音声・楽音復号化装置105について、図7の詳細ブロック図を用いて説明する。

[0138] 形状コードブック204、利得コードブック205は、それぞれ図2で示すものと同様である。

[0139] ベクトル復号化部701は、伝送路103を介して伝送される符号化情報102を入力とし、符号化情報である $code_index_{MIN}$ と $gain_index_{MIN}$ とを用いて、形状コードブック204からコードベクトル $codek^{code_index_{MIN}}$ ($k=0, \Lambda, N-1$)を読み込み、また、利得コードブック205から利得コード $gain^{gain_index_{MIN}}$ を読み込む。次に、ベクトル復号化部701は、 $gain^{gain_index_{MIN}}$ と $codek^{code_index_{MIN}}$ ($k=0, \Lambda, N-1$)とを乗算し、乗算した結果得られる $gain^{gain_index_{MIN}} \times codek^{code_index_{MIN}}$ ($k=0, \Lambda, N-1$)を復号化MDCT係数として直交変換処理部702に出力する。

[0140] 直交変換処理部702は、バッファ buf_k' を内部に有し、式(38)により初期化する。

[0141] [数38]

$$buf_k' = 0 \quad (k = 0, \dots, N - 1) \quad \dots (38)$$

[0142] 次に、MDCT係数復号化部701から出力される復号化MDCT係数 $gain^{gain_index_{MIN}}$

×code_k^{code_indexMIN}(k=0、Λ、N-1)を入力とし、式(39)により復号化信号 y_n を求める。
。

[0143] [数39]

$$y_n = \frac{2}{N} \sum_{k=0}^{2N-1} X'_k \cos \left[\frac{(2n+1+N)(2k+1)\pi}{4N} \right] \quad (n=0, \dots, N-1) \quad \dots (39)$$

[0144] ここで、 X'_k は、復号化MDCT係数 $gain^{gain_indexMIN} \times code_k^{code_indexMIN}$ (k=0、Λ、N-1)とバッファ buf'_k とを結合させたベクトルであり、式(40)により求める。

[0145] [数40]

$$X'_k = \begin{cases} buf'_k & (k=0, \dots, N-1) \\ gain^{gain_indexMIN} \cdot code_{k-N}^{code_indexMIN} & (k=N, \dots, 2N-1) \end{cases} \quad \dots (40)$$

[0146] 次に、式(41)によりバッファ buf'_k を更新する。

[0147] [数41]

$$buf'_k = gain^{gain_indexMIN} \cdot code_k^{code_indexMIN} \quad (k=0, \dots, N-1) \quad \dots (41)$$

[0148] 次に、復号化信号 y_n を出力信号106として出力する。

[0149] このように、入力信号のMDCT係数を求める直交変換処理部と、聴感マスキング特性値を求める聴感マスキング特性値算出部と、聴感マスキング特性値を利用したベクトル量子化を行うベクトル量子化部とを設け、聴感マスキング特性値とMDCT係数と量子化されたMDCT係数との相対的位置関係に応じてベクトル量子化の距離計算を行うことにより、聴感的に影響の大きい信号の劣化を抑える適切なコードベクトルを選択することができ、より高品質な出力信号を得ることができる。

[0150] なお、ベクトル量子化部202において、前記場合1から場合5の各距離計算に対し聴感重み付けフィルタを適用することにより量子化することも可能である。

[0151] なお、本実施の形態では、MDCT係数の符号化を行う場合について説明したが、フーリエ変換、離散コサイン変換(DCT)、及び直交鏡像フィルタ(QMF)等の直交変換を用いて、変換後の信号(周波数パラメータ)の符号化を行う場合についても本発明は適用することができ、本実施の形態と同様の作用・効果を得ることができる。

[0152] なお、本実施の形態では、ベクトル量子化により符号化を行う場合について説明したが、本発明は符号化方法に制限はなく、例えば、分割ベクトル量子化、多段階ベク

トル量子化により符号化を行ってもよい。

[0153] なお、音声・楽音符号化装置101を図16のフローチャートで示した手順をプログラムによりコンピュータで実行させてもよい。

[0154] 以上説明したように、入力信号から聴感マスキング特性値を算出し、入力信号のMDCT係数、符号化値、及び聴感マスキング特性値の相対的な位置関係を全て考慮し、人の聴感に適した距離計算法を適用することにより、聴感的に影響の大きい信号の劣化を抑える適切なコードベクトルを選択することができ、入力信号を低ビットレートで量子化した場合においても、より良好な復号化音声を得ることができる。

[0155] また、特許文献1では、図6の「場合5」のみ開示されているが、本発明においては、それらに加え、「場合2」、「場合3」、及び「場合4」に示されているように全ての組合せ関係においても、聴感マスキング特性値を考慮した距離計算法を採用することにより、入力信号のMDCT係数、符号化値及び聴感マスキング特性値の相対的な位置関係を全て考慮し、聴感に適した距離計算法を適用することで、入力信号を低ビットレートで量子化した場合においても、より良好な高品質な復号化音声を得ることができる。

[0156] また、本発明は、入力信号のMDCT係数または符号化値がこの聴感マスキング領域に存在した場合、また聴感マスキング領域を挟んで存在する場合、そのまま距離計算を行い、ベクトル量子化を行うと、実際の聴感が異なって聞こえるということに基づいたもので、ベクトル量子化の際の距離計算の方法を変えることにより、より自然な聴感を与えることができる。

[0157] (実施の形態2)

本発明の実施の形態2では、実施の形態1で説明した聴感マスキング特性値を用いたベクトル量子化をスケーラブル符号化に適用した例について説明する。

[0158] 以下、本実施の形態では、基本レイヤと拡張レイヤとで構成される二階層の音声符号化／復号化方法において拡張レイヤで聴感マスキング特性値を利用したベクトル量子化を行う場合について説明する。

[0159] スケーラブル音声符号化方法とは、周波数特性に基づき複数の階層(レイヤ)に音声信号を分解し符号化する方法である。具体的には、下位レイヤの入力信号と下位

レイヤの出力信号との差である残差信号を利用して各レイヤの信号を算出する。復号側ではこれら各レイヤの信号を加算し音声信号を復号する。この仕組みにより、音質を柔軟に制御できるほか、ノイズに強い音声信号の転送が可能となる。

- [0160] なお、本実施の形態では、基本レイヤがCELPタイプの音声符号化／復号化を行う場合を例にして説明する。
- [0161] 図8は、本発明の実施の形態2に係るMDCT係数ベクトル量子化方法を利用した符号化装置及び復号化装置の構成を示すブロック図である。なお、図8において、基本レイヤ符号化部801、基本レイヤ復号化部803及び拡張レイヤ符号化部805により符号化装置が構成され、基本レイヤ復号化部808、拡張レイヤ復号化部810及び加算部812により復号化装置が構成される。
- [0162] 基本レイヤ符号化部801は、入力信号800をCELPタイプの音声符号化方法を用いて符号化し、基本レイヤ符号化情報802を算出する共に、それを基本レイヤ復号化部803及び伝送路807を介して基本レイヤ復号化部808に出力する。
- [0163] 基本レイヤ復号化部803は、CELPタイプの音声復号化方法を用いて基本レイヤ符号化情報802を復号化し、基本レイヤ復号化信号804を算出すると共に、それを拡張レイヤ符号化部805に出力する。
- [0164] 拡張レイヤ符号化部805は、基本レイヤ復号化部803より出力される基本レイヤ復号化信号804と、入力信号800とを入力し、聴感マスキング特性値を利用したベクトル量子化により、入力信号800と基本レイヤ復号化信号804との残差信号を符号化し、符号化によって求められる拡張レイヤ符号化情報806を、伝送路807を介して拡張レイヤ復号化部810に出力する。拡張レイヤ符号化部805についての詳細は後述する。
- [0165] 基本レイヤ復号化部808は、CELPタイプの音声復号化方法を用いて基本レイヤ符号化情報802を復号化し、復号化によって求められる基本レイヤ復号化信号809を加算部812に出力する。
- [0166] 拡張レイヤ復号化部810は、拡張レイヤ符号化情報806を復号化し、復号化によって求められる拡張レイヤ復号化信号811を加算部812に出力する。
- [0167] 加算部812は、基本レイヤ復号化部808から出力された基本レイヤ復号化信号80

9と拡張レイヤ復号化部810から出力された拡張レイヤ復号化信号811とを加算し、加算結果である音声・楽音信号を出力信号813として出力する。

[0168] 次に、基本レイヤ符号化部801について図9のブロック図を用いて説明する。

[0169] 基本レイヤ符号化部801の入力信号800は、前処理部901に入力される。前処理部901は、DC成分を取り除くハイパスフィルタ処理や後続する符号化処理の性能改善につながるような波形整形処理やプリアンファシス処理を行い、これらの処理後の信号(X_{in})をLPC分析部902および加算部905に出力する。

[0170] LPC分析部902は、 X_{in} を用いて線形予測分析を行い、分析結果(線形予測係数)をLPC量子化部903へ出力する。LPC量子化部903は、LPC分析部902から出力された線形予測係数(LPC)の量子化処理を行い、量子化LPCを合成フィルタ904へ出力するとともに量子化LPCを表す符号(L)を多重化部914へ出力する。

[0171] 合成フィルタ904は、量子化LPCに基づくフィルタ係数により、後述する加算部911から出力される駆動音源に対してフィルタ合成を行うことにより合成信号を生成し、合成信号を加算部905へ出力する。

[0172] 加算部905は、合成信号の極性を反転させて X_{in} に加算することにより誤差信号を算出し、誤差信号を聴覚重み付け部912へ出力する。

[0173] 適応音源符号帳906は、過去に加算部911によって出力された駆動音源をバッファに記憶しており、パラメータ決定部913から出力された信号により特定される過去の駆動音源から1フレーム分のサンプルを適応音源ベクトルとして切り出して乗算部909へ出力する。

[0174] 量子化利得生成部907は、パラメータ決定部913から出力された信号によって特定される量子化適応音源利得と量子化固定音源利得とをそれぞれ乗算部909と乗算部910へ出力する。

[0175] 固定音源符号帳908は、パラメータ決定部913から出力された信号によって特定される形状を有するパルス音源ベクトルに拡散ベクトルを乗算して得られた固定音源ベクトルを乗算部910へ出力する。

[0176] 乗算部909は、量子化利得生成部907から出力された量子化適応音源利得を、適応音源符号帳906から出力された適応音源ベクトルに乗じて、加算部911へ出力す

る。乗算部910は、量子化利得生成部907から出力された量子化固定音源利得を、固定音源符号帳908から出力された固定音源ベクトルに乗じて、加算部911へ出力する。

[0177] 加算部911は、利得乗算後の適応音源ベクトルと固定音源ベクトルとをそれぞれ乗算部909と乗算部910とから入力し、これらをベクトル加算し、加算結果である駆動音源を合成フィルタ904および適応音源符号帳906へ出力する。なお、適応音源符号帳906に入力された駆動音源は、バッファに記憶される。

[0178] 聴覚重み付け部912は、加算部905から出力された誤差信号に対して聴覚的な重み付けをおこない符号化歪みとしてパラメータ決定部913へ出力する。

[0179] パラメータ決定部913は、聴覚重み付け部912から出力された符号化歪みを最小とする適応音源ベクトル、固定音源ベクトル及び量子化利得を、各々適応音源符号帳906、固定音源符号帳908及び量子化利得生成部907から選択し、選択結果を示す適応音源ベクトル符号(A)、音源利得符号(G)及び固定音源ベクトル符号(F)を多重化部914に出力する。

[0180] 多重化部914は、LPC量子化部903から量子化LPCを表す符号(L)を入力し、パラメータ決定部913から適応音源ベクトルを表す符号(A)、固定音源ベクトルを表す符号(F)および量子化利得を表す符号(G)を入力し、これらの情報を多重化して基本レイヤ符号化情報802として出力する。

[0181] 次に、基本レイヤ復号化部803(808)について図10を用いて説明する。

[0182] 図10において、基本レイヤ復号化部803(808)に入力された基本レイヤ符号化情報802は、多重化分離部1001によって個々の符号(L、A、G、F)に分離される。分離されたLPC符号(L)はLPC復号化部1002に出力され、分離された適応音源ベクトル符号(A)は適応音源符号帳1005に出力され、分離された音源利得符号(G)は量子化利得生成部1006に出力され、分離された固定音源ベクトル符号(F)は固定音源符号帳1007へ出力される。

[0183] LPC復号化部1002は、多重化分離部1001から出力された符号(L)から量子化LPCを復号化し、合成フィルタ1003に出力する。

[0184] 適応音源符号帳1005は、多重化分離部1001から出力された符号(A)で指定さ

れる過去の駆動音源から1フレーム分のサンプルを適応音源ベクトルとして取り出して乗算部1008へ出力する。

[0185] 量子化利得生成部1006は、多重化分離部1001から出力された音源利得符号(G)で指定される量子化適応音源利得と量子化固定音源利得を復号化し乗算部1008及び乗算部1009へ出力する。

[0186] 固定音源符号帳1007は、多重化分離部1001から出力された符号(F)で指定される固定音源ベクトルを生成し、乗算部1009へ出力する。

[0187] 乗算部1008は、適応音源ベクトルに量子化適応音源利得を乗算して、加算部1010へ出力する。乗算部1009は、固定音源ベクトルに量子化固定音源利得を乗算して、加算部1010へ出力する。

[0188] 加算部1010は、乗算部1008、乗算部1009から出力された利得乗算後の適応音源ベクトルと固定音源ベクトルの加算を行い、駆動音源を生成し、これを合成フィルタ1003及び適応音源符号帳1005に出力する。

[0189] 合成フィルタ1003は、LPC復号化部1002によって復号化されたフィルタ係数を用いて、加算部1010から出力された駆動音源のフィルタ合成を行い、合成した信号を後処理部1004へ出力する。

[0190] 後処理部1004は、合成フィルタ1003から出力された信号に対して、ホルマント強調やピッチ強調といったような音声の主観的な品質を改善する処理や、定常雑音の主観的な品質を改善する処理などを施し、基本レイヤ復号化信号804(810)として出力する。

[0191] 次に、拡張レイヤ符号化部805について図11を用いて説明する。

[0192] 図11の拡張レイヤ符号化部805は、図2と比較して、直交変換処理部1103への入力信号が基本レイヤ復号化信号804と入力信号800との差分信号1102が入力される以外は同様であり、聴感マスキング特性値算出部203には図2と同一符号を付して説明を省略する。

[0193] 拡張レイヤ符号化部805は、実施の形態1の符号化部101と同様に、入力信号800をNサンプルずつ区切り(Nは自然数)、Nサンプルを1フレームとしてフレーム毎に符号化を行う。ここで、符号化の対象となる入力信号800を x_n ($n=0, 1, \dots, N-1$)と表

すこととする。

[0194] 入力信号 x_n 800は、聴感マスキング特性値算出部203、及び加算部1101に入力される。また、基本レイヤ復号化部803から出力される基本レイヤ復号化信号804は、加算部1101、及び直交変換処理部1103に入力される。

[0195] 加算部1101は、式(42)により残差信号 $xresid_n$ ($n=0, \Lambda, N-1$)を求め、求めた残差信号 $xresid_n$ 1102を直交変換処理部1103に出力する。

[0196] [数42]

$$xresid_n = x_n - xbase_n \quad (n=0, \dots, N-1) \quad \dots (42)$$

[0197] ここで、 $xbase_n$ ($n=0, \Lambda, N-1$)は基本レイヤ復号化信号804である。次に、直交変換処理部1103の処理について説明する。

[0198] 直交変換処理部1103は、基本レイヤ復号化信号 $xbase_n$ 804の処理時に使用するバッファ $bufbase_n$ ($n=0, \Lambda, N-1$)と、残差信号 $xresid_n$ 1102の処理時に使用するバッファ $bufresid_n$ ($n=0, \Lambda, N-1$)を内部に有し、式(43)及び式(44)によってそれぞれ初期化する。

[0199] [数43]

$$bufbase_n = 0 \quad (n=0, \dots, N-1) \quad \dots (43)$$

[0200] [数44]

$$bufresid_n = 0 \quad (n=0, \dots, N-1) \quad \dots (44)$$

[0201] 次に、直交変換処理部1103は、基本レイヤ復号化信号 $xbase_n$ 804と残差信号 $xresid_n$ 1102とを修正離散コサイン変換(MDCT)することにより、基本レイヤ直交変換係数 $xbase_k$ 1104と残差直交変換係数 $Xresid_k$ 1105とをそれぞれ求める。ここで、基本レイヤ直交変換係数 $xbase_k$ 1104は式(45)により求める。

[0202] [数45]

$$Xbase_k = \frac{2}{N} \sum_{n=0}^{2N-1} xbase'_n \cos \left[\frac{(2n+1+N)(2k+1)\pi}{4N} \right] \quad (k=0, \dots, N-1) \quad \dots (45)$$

[0203] ここで、 $xbase'_n$ は基本レイヤ復号化信号 $xbase_n$ 804とバッファ $bufbase_n$ とを結合したベクトルであり、直交変換処理部1103は、式(46)により $xbase'_n$ を求める。また、 k は1フレームにおける各サンプルのインデックスである。

[0204] [数46]

$$xbase'_n = \begin{cases} bufbase_n & (n = 0, \dots, N-1) \\ xbase_{n-N} & (n = N, \dots, 2N-1) \end{cases} \quad \dots (46)$$

[0205] 次に、直交変換処理部1103は、式(47)によりバッファbufbase_nを更新する。

[0206] [数47]

$$bufbase_n = xbase_n \quad (n = 0, \dots, N-1) \quad \dots (47)$$

[0207] また、直交変換処理部1103は、式(48)により残差直交変換係数Xresid_k 1105を求める。

[0208] [数48]

$$Xresid_k = \frac{2}{N} \sum_{n=0}^{2N-1} xresid'_n \cos \left[\frac{(2n+1+N)(2k+1)\pi}{4N} \right] \quad (k = 0, \dots, N-1) \quad \dots (48)$$

[0209] ここで、xresid'_nは残差信号xresid_n 1102とバッファbufresid_nとを結合したベクトルであり、直交変換処理部1103は、式(49)によりxresid_n'を求める。また、kは1フレームにおける各サンプルのインデックスである。

[0210] [数49]

$$xresid'_n = \begin{cases} bufresid_n & (n = 0, \dots, N-1) \\ xresid_{n-N} & (n = N, \dots, 2N-1) \end{cases} \quad \dots (49)$$

[0211] 次に、直交変換処理部1103は、式(50)によりバッファbufresid_nを更新する。

[0212] [数50]

$$bufresid_n = xresid_n \quad (n = 0, \dots, N-1) \quad \dots (50)$$

[0213] 次に、直交変換処理部1103は、基本レイヤ直交変換係数Xbase_k 1104と残差直交変換係数Xresid_k 1105とをベクトル量子化部1106に出力する。

[0214] ベクトル量子化部1106は、直交変換処理部1103から基本レイヤ直交変換係数Xbase_k 1104と残差直交変換係数Xresid_k 1105と、聴感マスキング特性値算出部203から聴感マスキング特性値M_k 1107とを入力し、形状コードブック1108と利得コードブック1109とを用いて、聴感マスキング特性値を利用したベクトル量子化により残差直交変換係数Xresid_k 1105の符号化を行い、符号化により得られる拡張レイヤ符号化情報806を出力する。

[0215] ここで、形状コードブック1108は、予め作成された N_e 種類の N 次元コードベクトル coderesid_k^e ($e=0, \Lambda, N_e-1, k=0, \Lambda, N-1$)から構成され、前記ベクトル量子化部1103において残差直交変換係数 $X\text{resid}_k$ 1105をベクトル量子化する際に用いられる。

[0216] また、利得コードブック1109は、予め作成された N_f 種類の残差利得コード gainresid_f ($f=0, \Lambda, N_f-1$)から構成され、前記ベクトル量子化部1106において残差直交変換係数 $X\text{resid}_k$ 1105をベクトル量子化する際に用いられる。

[0217] 次に、ベクトル量子化部1106の処理について、図12を用いて詳細に説明する。ステップ1201では、形状コードブック1108におけるコードベクトルインデックス e に0を代入し、最小誤差 Dist_{MIN} を十分大きな値を代入し、初期化する。

[0218] ステップ1202では、図11の形状コードブック1108から N 次元のコードベクトル coderesid_k^e ($k=0, \Lambda, N-1$)を読み込む。

[0219] ステップ1203では、直交変換処理部1103から出力された残差直交変換係数 $X\text{resid}_k$ を入力し、ステップ1202で読み込んだコードベクトル coderesid_k^e ($k=0, \Lambda, N-1$)の利得 Gainresid を式(51)により求める。

[0220] [数51]

$$\text{Gainresid} = \sum_{k=0}^{N-1} X\text{resid}_k \cdot \text{coderesid}_k^e \bigg/ \sum_{k=0}^{N-1} \text{coderesid}_k^{e^2} \quad \dots (51)$$

[0221] ステップ1204では、ステップ1205の実行回数を表す $\text{calc_count}_{\text{resid}}$ に0を代入する。

[0222] ステップ1205では、聴感マスキング特性値算出部203から出力された聴感マスキング特性値 M_k を入力とし、式(52)により一時利得 temp2_k ($k=0, \Lambda, N-1$)を求める。

[0223] [数52]

$$\text{temp2}_k = \begin{cases} \text{coderesid}_k^e & \left(\left| \text{coderesid}_k^e \cdot \text{Gainresid} + X\text{base}_k \right| \geq M_k \right) \\ 0 & \left(\left| \text{coderesid}_k^e \cdot \text{Gainresid} + X\text{base}_k \right| < M_k \right) \end{cases} \quad (k=0, \dots, N-1)$$

\dots (52)

[0224] なお、式(52)において、 k が $\left| \text{coderesid}_k^e \cdot \text{Gainresid} + X\text{base}_k \right| \geq M_k$ の条件を

満たす場合、一時利得 $temp2_k$ には $coderesid_k^e$ が代入され、 k が $|coderesid_k^e \cdot Gainresid + Xbase_k| < M_k$ の条件を満たす場合、 $temp2_k$ には0が代入される。また、 k は1フレームにおける各サンプルのインデックスである。

[0225] 次に、ステップ1205では、式(53)により利得Gainresidを求める。

[0226] [数53]

$$Gainresid = \sum_{k=0}^{N-1} Xresid_k \cdot temp2_k / \sum_{k=0}^{N-1} temp2_k^2 \quad (k=0, \dots, N-1) \quad \dots (53)$$

[0227] ここで、全ての k において一時利得 $temp2_k$ が0の場合には利得Gainresidに0を代入する。また、式(54)により、利得Gainresidとコードベクトル $coderesid_k^e$ から残差符号化値 $Rresid_k$ を求める。

[0228] [数54]

$$Rresid_k = Gainresid \cdot coderesid_k^e \quad (k=0, \dots, N-1) \quad \dots (54)$$

[0229] また、式(55)により、残差符号化値 $Rresid_k$ と基本レイヤ直交変換係数 $Xbase_k$ から加算符号化値 $Rplus_k$ を求める。

[0230] [数55]

$$Rplus_k = Rresid_k + Xbase_k \quad (k=0, \dots, N-1) \quad \dots (55)$$

[0231] ステップ1206では、 $calc_count_{resid}$ に1を足し加える。

[0232] ステップ1207では、 $calc_count_{resid}$ と予め定められた非負の整数 $Nresid_c$ とを比較し、 $calc_count_{resid}$ が $Nresid_c$ より小さい値である場合はステップ1205に戻り、 $calc_count_{resid}$ が $Nresid_c$ 以上である場合はステップ1208に進む。

[0233] ステップ1208では、累積誤差Distresidに0を代入し、また、 k に0を代入する。また、ステップ1208では、式(56)により加算MDCT係数 $Xplus_k$ を求める。

[0234] [数56]

$$Xplus_k = Xbase_k + Xresid_k \quad (k=0, \dots, N-1) \quad \dots (56)$$

[0235] 次に、ステップ1209、1211、1212、及び1214において、聴感マスキング特性値 $Mk1107$ と加算符号化値 $Rplus_k$ と加算MDCT係数 $Xplus_k$ との相対的な位置関係について場合分けを行い、場合分けの結果に応じてそれぞれステップ1210、1213、1215、及び1216で距離計算する。この相対的な位置関係による場合分けを図13

に示す。図13において、白い丸記号(○)は加算MDCT係数 X_{plus_k} を意味し、黒い丸記号(●)は R_{plus_k} を意味するものである。図13における考え方は、実施の形態1の図6で説明した考え方と同様である。

[0236] ステップ1209では、聴感マスキング特性値 M_k と加算符号化値 R_{plus_k} と加算MDCT係数 X_{plus_k} との相対的な位置関係が図13における「場合1」に該当するかどうかを式(57)の条件式により判定する。

[0237] [数57]

$$(|X_{plus_k}| \geq M_k) \text{ and } (|R_{plus_k}| \geq M_k) \text{ and } (X_{plus_k} \cdot R_{plus_k} \geq 0) \quad \dots (57)$$

[0238] 式(57)は、加算MDCT係数 X_{plus_k} の絶対値と加算符号化値 R_{plus_k} の絶対値とが共に聴感マスキング特性値 M_k 以上であり、かつ、加算MDCT係数 X_{plus_k} と加算符号化値 R_{plus_k} とが同符号である場合を意味する。聴感マスキング特性値 M_k と加算MDCT係数 X_{plus_k} と加算符号化値 R_{plus_k} とが式(57)の条件式を満たした場合は、ステップ1210に進み、式(57)の条件式を満たさない場合は、ステップ1211に進む。

[0239] ステップ1210では、式(58)により R_{plus_k} と加算MDCT係数 X_{plus_k} との誤差 $Distr_{resid}$ を求め、累積誤差 $Dist_{resid}$ に誤差 $Distr_{resid}$ を加算し、ステップ1217に進む。

[0240] [数58]

$$\begin{aligned} Dist_{resid}_1 &= Dresid_{11} \\ &= |X_{resid_k} - R_{resid_k}| \quad \dots (58) \end{aligned}$$

[0241] ステップ1211では、聴感マスキング特性値 M_k と加算符号化値 R_{plus_k} と加算MDCT係数 X_{plus_k} との相対的な位置関係が図13における「場合5」に該当するかどうかを式(59)の条件式により判定する。

[0242] [数59]

$$(|X_{plus_k}| < M_k) \text{ and } (|R_{plus_k}| < M_k) \quad \dots (59)$$

[0243] 式(59)は、加算MDCT係数 X_{plus_k} の絶対値と加算符号化値 R_{plus_k} の絶対値とが共に聴感マスキング特性値 M_k 未満である場合を意味する。聴感マスキング特性値 M_k と加算符号化値 R_{plus_k} と加算MDCT係数 X_{plus_k} が式(59)の条件式を満たす場

合、加算符号化値 $Rplus_k$ と加算MDCT係数 $Xplus_k$ との誤差は0とし、累積誤差 $Dist_{resid}$ には何も加算せずにステップ1217に進む。聴感マスキング特性値 M_k と加算符号化値 $Rplus_k$ と加算MDCT係数 $Xplus_k$ が式(59)の条件式を満たさない場合は、ステップ1212に進む。

- [0244] ステップ1212では、聴感マスキング特性値 M_k と加算符号化値 $Rplus_k$ と加算MDCT係数 $Xplus_k$ との相対的な位置関係が図13における「場合2」に該当するかどうかを式(60)の条件式により判定する。

- [0245] [数60]

$$(|Xplus_k| \geq M_k) \text{ and } (|Rplus_k| \geq M_k) \text{ and } (Xplus_k \cdot Rplus_k < 0) \quad \dots (60)$$

- [0246] 式(60)は、加算MDCT係数 $Xplus_k$ の絶対値と加算符号化値 $Rplus_k$ の絶対値とが共に聴感マスキング特性値 M_k 以上であり、かつ、加算MDCT係数 $Xplus_k$ と加算符号化値 $Rplus_k$ とが異符号である場合を意味する。聴感マスキング特性値 M_k と加算MDCT係数 $Xplus_k$ と加算符号化値 $Rplus_k$ とが式(60)の条件式を満たした場合は、ステップ1213に進み、式(60)の条件式を満たさない場合は、ステップ1214に進む。

- [0247] ステップ1213では、式(61)により加算符号化値 $Rplus_k$ と加算MDCT係数 $Xplus_k$ との誤差 $Dist_{resid}_2$ を求め、累積誤差 $Dist_{resid}$ に誤差 $Dist_{resid}_2$ を加算し、ステップ1217に進む。

- [0248] [数61]

$$Dist_{resid}_2 = Dresid_{21} + Dresid_{22} + \beta_{resid} * Dresid_{23} \quad \dots (61)$$

- [0249] ここで、 β_{resid} は、加算MDCT係数 $Xplus_k$ 、加算符号化値 $Rplus_k$ 及び聴感マスキング特性値 M_k に応じて適宜設定される値であり、1以下の値が適当である。また、 $Dresid_{21}$ 、 $Dresid_{22}$ 及び $Dresid_{23}$ は、それぞれ式(62)、式(63)及び式(64)により求められる。

- [0250] [数62]

$$Dresid_{21} = |Xplus_k| - M_k \quad \dots (62)$$

- [0251] [数63]

$$Dresid_{22} = |Rplus_k| - M_k \quad \dots (63)$$

[0252] [数64]

$$Dresid_{23} = M_k \cdot 2 \quad \dots (64)$$

[0253] ステップ1214では、聴感マスキング特性値 M_k と加算符号化値 $Rplus_k$ と加算MDCT係数 $Xplus_k$ との相対的な位置関係が図13における「場合3」に該当するかどうかを式(65)の条件式により判定する。

[0254] [数65]

$$(|Xplus_k| \geq M_k) \text{ and } (|Rplus_k| < M_k) \quad \dots (65)$$

[0255] 式(65)は、加算MDCT係数 $Xplus_k$ の絶対値が聴感マスキング特性値 M_k 以上であり、かつ、加算符号化値 $Rplus_k$ が聴感マスキング特性値 M_k 未満である場合を意味する。聴感マスキング特性値 M_k と加算MDCT係数 $Xplus_k$ と加算符号化値 $Rplus_k$ とが式(65)の条件式を満たした場合は、ステップ1215に進み、式(65)の条件式を満たさない場合は、ステップ1216に進む。

[0256] ステップ1215では、式(66)により加算符号化値 $Rplus_k$ と加算MDCT係数 $Xplus_k$ との誤差 $Distresid_3$ を求め、累積誤差 $Distresid$ に誤差 $Distresid_3$ を加算し、ステップ1217に進む。

[0257] [数66]

$$\begin{aligned} Distresid_3 &= Dresid_{31} \\ &= |Xplus_k| - M_k \end{aligned} \quad \dots (66)$$

[0258] ステップ1216では、聴感マスキング特性値 M_k と加算符号化値 $Rplus_k$ と加算MDCT係数 $Xplus_k$ との相対的な位置関係が図13における「場合4」に該当し、式(67)の条件式を満たす。

[0259] [数67]

$$(|Xplus_k| < M_k) \text{ and } (|Rplus_k| \geq M_k) \quad \dots (67)$$

[0260] 式(67)は、加算MDCT係数 $Xplus_k$ の絶対値が聴感マスキング特性値 M_k 未満であり、かつ、加算符号化値 $Rplus_k$ が聴感マスキング特性値 M_k 以上である場合を意味する。この時、ステップ1216は、式(68)により加算符号化値 $Rplus_k$ と加算MDCT係数 $Xplus_k$ との誤差 $Distresid_4$ を求め、累積誤差 $Distresid$ に誤差 $Distresid_4$ を加算し、ステップ1217に進む。

[0261] [数68]

$$\begin{aligned} \text{Distresid}_k &= \text{Dresid}_{41} \quad \dots (68) \\ &= |Rplus_k| - M_k \end{aligned}$$

[0262] ステップ1217では、kに1を足し加える。

[0263] ステップ1218では、Nとkを比較し、kがNより小さい値の場合は、ステップ1209に戻る。kがN以上である場合は、ステップ1219に進む。

[0264] ステップ1219では、累積誤差Distresidと最小誤差Distresid_{MIN}とを比較し、累積誤差Distresidが最小誤差Distresid_{MIN}より小さい値の場合は、ステップ1220に進み、累積誤差Distresidが最小誤差Distresid_{MIN}以上である場合は、ステップ1221に進む。

[0265] ステップ1220では、最小誤差Distresid_{MIN}に累積誤差Distresidを代入し、gainresid_index_{MIN}にeを代入し、誤差最小利得Distresid_{MIN}に利得Distresidを代入し、ステップ1221に進む。

[0266] ステップ1221では、eに1を足し加える。

[0267] ステップ1222では、コードベクトルの総数N_eとeとを比較し、eがN_eより小さい値の場合は、ステップ1202に戻る。eがN_e以上である場合は、ステップ1223に進む。

[0268] ステップ1223では、図11の利得コードブック1109からN_f種類の残差利得コードgainresid^f(f=0、Λ、N_f-1)を読み込み、全てのfに対して式(69)により量子化残差利得誤差gainresiderr^f(f=0、Λ、N_f-1)を求める。

[0269] [数69]

$$\text{gainresiderr}^f = | \text{Gainresid}_{\text{MIN}} - \text{gainresid}^f | \quad (f = 0, \dots, N_f - 1) \quad \dots (69)$$

[0270] 次に、ステップ1223では、量子化残差利得誤差gainresiderr^f(f=0、Λ、N_f-1)を最小とするfを求め、求めたfをgainresid_index_{MIN}に代入する。

[0271] ステップ1224では、累積誤差Distresidが最小となるコードベクトルのインデックスであるgainresid_index_{MIN}、及びステップ1223で求めたgainresid_index_{MIN}を拡張レイヤ符号化情報806として、伝送路807に出力し、処理を終了する。

[0272] 次に、拡張レイヤ復号化部810について、図14のブロック図を用いて説明する。形状コードブック1403は、形状コードブック1108と同様に、N_e種類のN次元コードベク

トル gainresid_k^e ($e=0, \Lambda, N_e-1, k=0, \Lambda, N-1$)から構成される。また利得コードブック1404は、利得コードブック1109と同様に、 N_f 種類の残差利得コード gainresid_f^e ($f=0, \Lambda, N_f-1$)から構成される。

[0273] ベクトル復号化部1401は、伝送路807を介して伝送される拡張レイヤ符号化情報806を入力とし、符号化情報である $\text{gainresid_index}_{\text{MIN}}$ と $\text{gainresid_index}_{\text{MIN}}$ とを用いて、形状コードブック1403からコードベクトル $\text{coderesid}_k^{\text{coderesid_indexMIN}}$ ($k=0, \Lambda, N-1$)を読み込み、また利得コードブック1404からコード $\text{gainresid}^{\text{gainresid_indexMIN}}$ を読み込む。次に、ベクトル復号化部1401は、 $\text{gainresid}^{\text{gainresid_indexMIN}}$ と $\text{coderesid}_k^{\text{coderesid_indexMIN}}$ ($k=0, \Lambda, N-1$)を乗算し、乗算した結果得られる $\text{gainresid}^{\text{gainresid_indexMIN}} \cdot \text{coderesid}_k^{\text{coderesid_indexMIN}}$ ($k=0, \Lambda, N-1$)を復号化残差直交変換係数として残差直交変換処理部1402に出力する。

[0274] 次に、残差直交変換処理部1402の処理について説明する。

[0275] 残差直交変換処理部1402は、バッファ $\text{bufresid}_k'$ を内部に有し、式(70)により初期化される。

[0276] [数70]

$$\text{bufresid}'_k = 0 \quad (k=0, \dots, N-1) \quad \dots (70)$$

[0277] 残差直交変換係数復号化部1401から出力される復号化残差直交変換係数 $\text{gainresid}^{\text{gainresid_indexMIN}} \cdot \text{coderesid}_k^{\text{coderesid_indexMIN}}$ ($k=0, \Lambda, N-1$)を入力して、式(71)により拡張レイヤ復号化信号 yresid_n 811を求める。

[0278] [数71]

$$\text{yresid}_n = \frac{2}{N} \sum_{k=0}^{2N-1} X\text{resid}'_k \cos \left[\frac{(2n+1+N)(2k+1)\pi}{4N} \right] \quad (n=0, \dots, N-1) \quad \dots (71)$$

[0279] ここで、 $X\text{resid}'_k$ は復号化残差直交変換係数 $\text{gainresid}^{\text{gainresid_indexMIN}} \cdot \text{coderesid}_k^{\text{coderesid_indexMIN}}$ ($k=0, \Lambda, N-1$)とバッファ $\text{bufresid}'_k$ とを結合させたベクトルであり、式(72)により求める。

[0280] [数72]

$$X\text{resid}'_k = \begin{cases} \text{bufresid}'_k & (k=0, \dots, N-1) \\ \text{gainresid}^{\text{gainresid_indexMIN}} \cdot \text{coderesid}_{k-N}^{\text{coderesid_indexMIN}} & (k=N, \dots, 2N-1) \end{cases} \quad \dots (72)$$

[0281] 次に、式(73)によりバッファ $\text{bufresid}'_k$ を更新する。

[0282] [数73]

$$\text{bufresid}'_k = \text{gainresid}^{\text{gainresid_index}_{\text{MAX}}} \cdot \text{coderesid}_k^{\text{coderesid_index}_{\text{MAX}}} \quad (k = 0, \dots, N-1) \quad \dots (73)$$

[0283] 次に、拡張レイヤ復号化信号 yresid_n 11を出力する。

[0284] なお、本発明はスケーラブル符号化の階層について制限はなく、三階層以上の階層的な音声符号化／復号化方法において上位レイヤで聴感マスキング特性値を利用したベクトル量子化を行う場合についても適用することができる。

[0285] なお、ベクトル量子化部1106において、前記場合1から場合5の各距離計算に対し聴感重み付けフィルタを適用することにより量子化してもよい。

[0286] なお、本実施の形態では、基本レイヤ符号化部／復号化部の音声符号化／復号化方法としてCELPタイプの音声符号化／復号化方法を例に挙げ説明したが、その他の音声符号化／復号化方法を用いてもよい。

[0287] なお、本実施の形態では、基本レイヤ符号化情報及び拡張レイヤ符号化情報を別々に送信する例を提示したが、各レイヤの符号化情報を多重化して送信し、復号側で多重化分離して各レイヤの符号化情報を復号するよう構成してもよい。

[0288] このように、スケーラブル符号化方式においても、本発明の聴感マスキング特性値を利用したベクトル量子化を適用することにより、聴感的に影響の大きい信号の劣化を抑える適切なコードベクトルを選択することができ、より高品質な出力信号を得ることができる。

[0289] (実施の形態3)

図15は、本発明の実施の形態3における上記実施の形態1、2で説明した符号化装置及び復号化装置を含む音声信号送信装置及び音声信号受信装置の構成を示すブロック図である。より具体的な応用としては、携帯電話、カーナビゲーションシステム等に適応可能である。

[0290] 図15において、入力装置1502は、音声信号1500をデジタル信号にA/D変換し音声・楽音符号化装置1503へ出力する。音声・楽音符号化装置1503は、図1に示した音声・楽音符号化装置101を実装し、入力装置1502から出力されたデジタル音声信号を符号化し、符号化情報をRF変調装置1504へ出力する。RF変調装置150

4は音声・楽音符号化装置1503から出力された音声符号化情報を電波等の伝播媒体に載せて送出するための信号に変換し送信アンテナ1505へ出力する。送信アンテナ1505はRF変調装置1504から出力された出力信号を電波(RF信号)として送出する。なお、図中のRF信号1506は送信アンテナ1505から送出された電波(RF信号)を表す。以上が音声信号送信装置の構成および動作である。

[0291] RF信号1507は受信アンテナ1508によって受信されRF復調装置1509へ出力される。なお、図中のRF信号1507は受信アンテナ1508に受信された電波を表し、伝播路において信号の減衰や雑音の重畳がなければRF信号1506と全く同じものになる。

[0292] RF復調装置1509は受信アンテナ1508から出力されたRF信号から音声符号化情報を復調し、音声・楽音復号化装置1510へ出力する。音声・楽音復号化装置1510は、図1に示した音声・楽音復号化装置105を実装し、RF復調装置1509から出力された音声符号化情報から音声信号を復号化し、出力装置1511は、復号されたデジタル音声信号をアナログ信号にD/A変換し、電氣的信号を空気の振動に変換し音波として人間の耳に聴こえるように出力する。

[0293] このように、音声信号送信装置及び音声信号受信装置においても、高品質な出力信号を得ることができる。

[0294] 本明細書は、2003年12月26日出願の特願2003-433160に基づくものである。この内容を全てここに含めておく。

産業上の利用可能性

[0295] 本発明は、聴感マスキング特性値を利用したベクトル量子化を適用することにより、聴感的に影響の大きい信号の劣化を抑える適切なコードベクトルを選択することができ、より高品質な出力信号を得ることができるという効果を有し、インターネット通信に代表されるパケット通信システムや、携帯電話、カーナビゲーションシステム等の移動通信システムの分野で、適応可能である。

請求の範囲

- [1] 音声・楽音信号を時間成分から周波数成分へ変換する直交変換処理手段と、前記音声・楽音信号から聴感マスキング特性値を求める聴感マスキング特性値算出手段と、前記聴感マスキング特性値に基づいて、予め設定されたコードブックから求めたコードベクトルと前記周波数成分との距離計算方法を変えてベクトル量子化を行うベクトル量子化手段と、を具備する音声・楽音符号化装置。
- [2] 音声・楽音信号を符号化して基本レイヤ符号化情報を生成する基本レイヤ符号化手段と、前記基本レイヤ符号化情報を復号化して基本レイヤ復号化信号を生成する基本レイヤ復号化手段と、前記音声・楽音信号と前記基本レイヤ復号化信号との差分信号を符号化して拡張レイヤ符号化情報を生成する拡張レイヤ符号化手段と、を具備する音声・楽音符号化装置において、前記拡張レイヤ符号化手段は、前記音声・楽音信号から聴感マスキング特性値を求める聴感マスキング特性値算出手段と、前記差分信号を時間成分から周波数成分へ変換する直交変換処理手段と、前記聴感マスキング特性値に基づいて、予め設定されたコードブックから求めたコードベクトルと前記周波数成分との距離計算方法を変えてベクトル量子化を行うベクトル量子化手段と、を具備する音声・楽音符号化装置。
- [3] 前記ベクトル量子化手段は、前記音声・楽音信号の周波数成分または前記コードベクトルのいずれか一方が、前記聴感マスキング特性値の示す聴感マスキング領域内にある場合に、前記聴感マスキング特性値に基づいて、前記音声・楽音信号の周波数成分と、前記コードベクトル間の距離計算方法を変えてベクトル量子化を行う請求項1記載の音声・楽音符号化装置。
- [4] 前記ベクトル量子化手段は、形状コードブックから求めたコードベクトル及び利得コードブックから求めたコードベクトルに基づきベクトル量子化を行う請求項1記載の音声・楽音信号符号化装置。
- [5] 前記直交変換処理手段は、修正離散コサイン変換(MDCT)、離散コサイン変換(DCT)、フーリエ変換または直交鏡像フィルタ(QMF)のいずれかにより前記音声・楽音信号を時間成分から周波数成分へ変換する請求項1記載の音声・楽音信号符号化装置。

- [6] さらに、少なくとも一つの拡張レイヤ符号化手段を具備し、前記拡張レイヤ符号化手段は、上位の拡張レイヤ符号化手段に対する入力信号と、前記上位の拡張レイヤ符号化手段が生成した拡張レイヤ符号化情報の復号化信号との差分を符号化して拡張レイヤ符号化情報を生成する請求項2記載の音声・楽音符号化装置。
- [7] 基本レイヤ符号化手段は、CELPタイプの音声・楽音信号符号化により入力信号を符号化する請求項2記載の音声・楽音信号符号化装置。
- [8] 音声・楽音信号を時間成分から周波数成分へ変換する直交変換処理ステップと、前記音声・楽音信号から聴感マスキング特性値を求める聴感マスキング特性値算出ステップと、前記聴感マスキング特性値に基づいて、予め設定されたコードブックから求めたコードベクトルと前記周波数成分との距離計算方法を変えてベクトル量子化を行うベクトル量子化ステップと、を具備する音声・楽音符号化方法。
- [9] コンピュータを、音声・楽音信号を時間成分から周波数成分へ変換する直交変換処理手段と、前記音声・楽音信号から聴感マスキング特性値を求める聴感マスキング特性値算出手段と、前記聴感マスキング特性値に基づいて、予め設定されたコードブックから求めたコードベクトルと前記周波数成分との距離計算方法を変えてベクトル量子化を行うベクトル量子化手段として機能させるための音声・楽音符号化プログラム。

補正書の請求の範囲

[2005年5月16日 (16. 05. 05) 国際事務局受理：出願当初の請求の範囲

1 - 9は取り下げられた；新しい請求の範囲

10 - 15が加えられた。(3頁)]

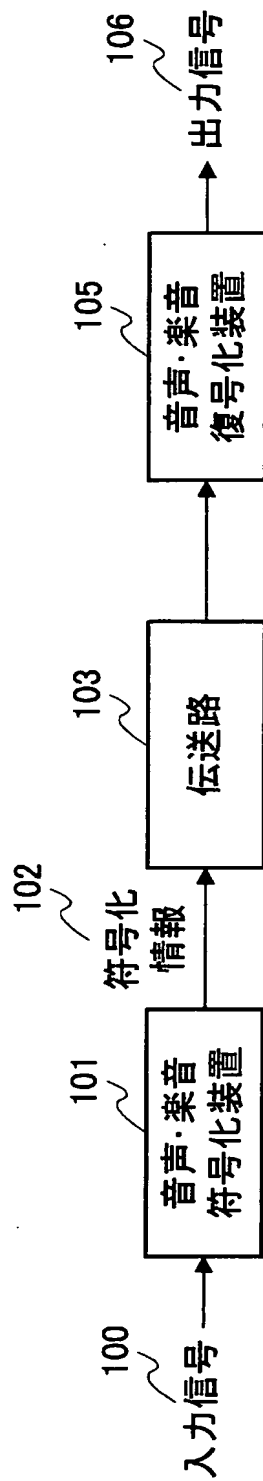
- [1] (削除)
- [2] (削除)
- [3] (削除)
- [4] (削除)
- [5] (削除)

- [6] (削除)
- [7] (削除)
- [8] (削除)
- [9] (削除)
- [10] (追加) 音声・楽音信号を時間成分から周波数成分へ変換する直交変換処理手段と、前記音声・楽音信号から聴感マスキング特性値を求める聴感マスキング特性値算出手段と、前記音声・楽音信号の周波数成分または前記コードベクトルのいずれか一方が前記聴感マスキング特性値の示す聴感マスキング領域内にある場合に、前記聴感マスキング特性値に基づいて前記音声・楽音信号の周波数成分と前記コードベクトル間の距離計算方法を変えてベクトル量子化を行うベクトル量子化手段と、を具備する音声・楽音符号化装置。
- [11] (追加) 音声・楽音信号を時間成分から周波数成分へ変換する直交変換処理手段と、前記音声・楽音信号から聴感マスキング特性値を求める聴感マスキング特性値算出手段と、前記音声・楽音信号の周波数成分と前記コードベクトルの符号が異なり、かつ、前記音声・楽音信号の周波数成分及び前記コードベクトルの符号が前記聴感マスキング特性値の示す聴感マスキング領域外にある場合に、前記聴感マスキング特性値に基づいて前記音声・楽音信号の周波数成分と前記コードベクトル間の距離計算方法を変えてベクトル量子化を行うベクトル量子化手段と、を具備する音声・楽音符号化装置。
- [12] (追加) 音声・楽音信号を時間成分から周波数成分へ変換する直交変換処理ステップと、前記音声・楽音信号から聴感マスキング特性値を求める聴感マスキング特性値算出ステップと、前記音声・楽音信号の周波数成分または前記コードベクトルのいずれか一方が前記聴感マスキング特性値の示す聴感マスキング領域内にある場合に、前記聴感マスキング特性値に基づいて前記音声・楽音信号の周波数成分と前記コードベクトル間の距離計算方法を変えてベクトル量子化を行うベクトル量子化ステップと、を具備する音声・楽音符号化方法。
- [13] (追加) 音声・楽音信号を時間成分から周波数成分へ変換する直交変換処理ステップと、前記音声・楽音信号から聴感マスキング特性値を求める聴感マスキング特性値算出ステップと、前記音声・楽音信号の周波数成分と前記コードベクトルの符号が異な

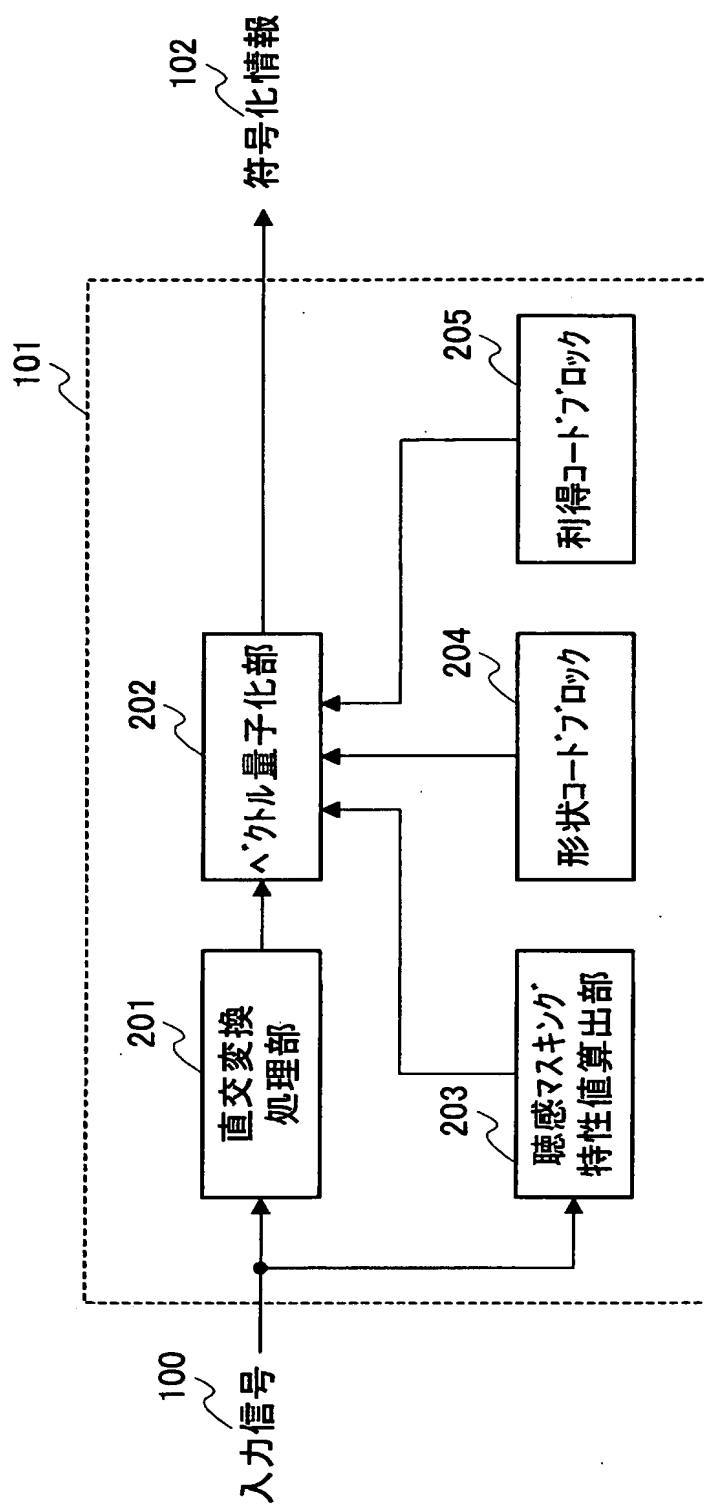
り、かつ、前記音声・楽音信号の周波数成分及び前記コードベクトルの符号が前記聴感マスキング特性値の示す聴感マスキング領域外にある場合に、前記聴感マスキング特性値に基づいて前記音声・楽音信号の周波数成分と前記コードベクトル間の距離計算方法を変えてベクトル量子化を行うベクトル量子化ステップと、を具備する音声・楽音符号化方法。

- [14] (追加) コンピュータを、音声・楽音信号を時間成分から周波数成分へ変換する直交変換処理手段と、前記音声・楽音信号から聴感マスキング特性値を求める聴感マスキング特性値算出手段と、前記音声・楽音信号の周波数成分または前記コードベクトルのいずれか一方が前記聴感マスキング特性値の示す聴感マスキング領域内にある場合に、前記聴感マスキング特性値に基づいて前記音声・楽音信号の周波数成分と前記コードベクトル間の距離計算方法を変えてベクトル量子化を行うベクトル量子化手段として機能させるための音声・楽音符号化プログラム。
- [15] (追加) コンピュータを、音声・楽音信号を時間成分から周波数成分へ変換する直交変換処理手段と、前記音声・楽音信号から聴感マスキング特性値を求める聴感マスキング特性値算出手段と、前記音声・楽音信号の周波数成分と前記コードベクトルの符号が異なり、かつ、前記音声・楽音信号の周波数成分及び前記コードベクトルの符号が前記聴感マスキング特性値の示す聴感マスキング領域外にある場合に、前記聴感マスキング特性値に基づいて前記音声・楽音信号の周波数成分と前記コードベクトル間の距離計算方法を変えてベクトル量子化を行うベクトル量子化手段として機能させるための音声・楽音符号化プログラム。

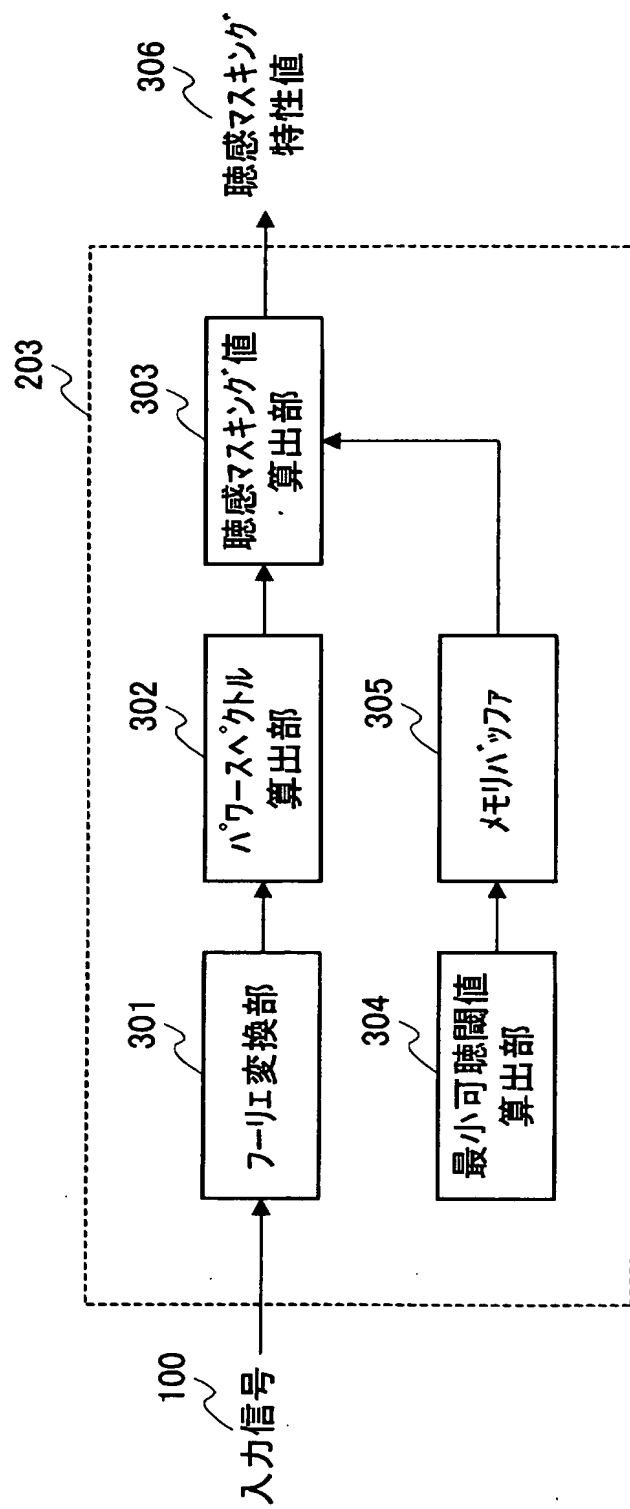
[図1]



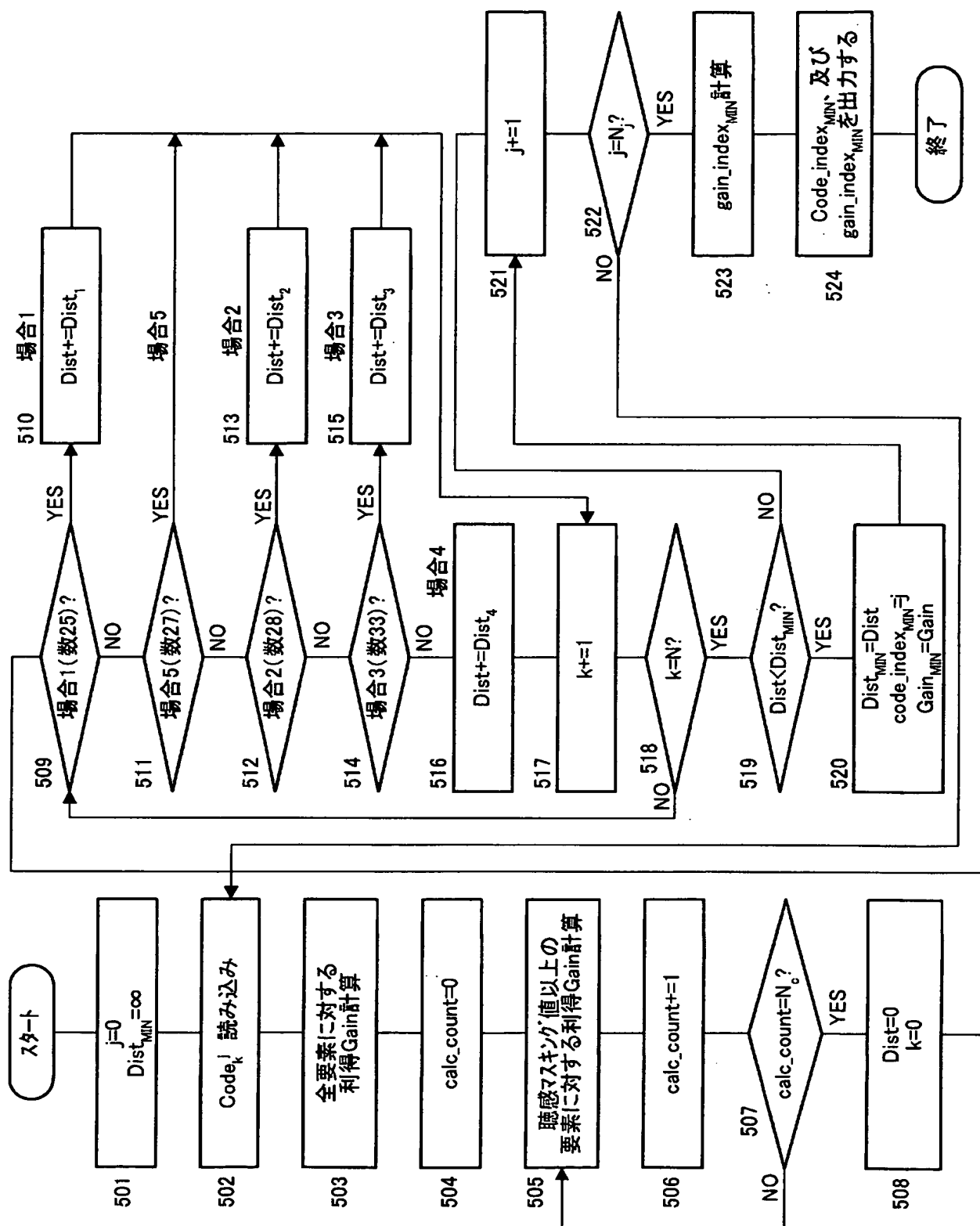
[図2]



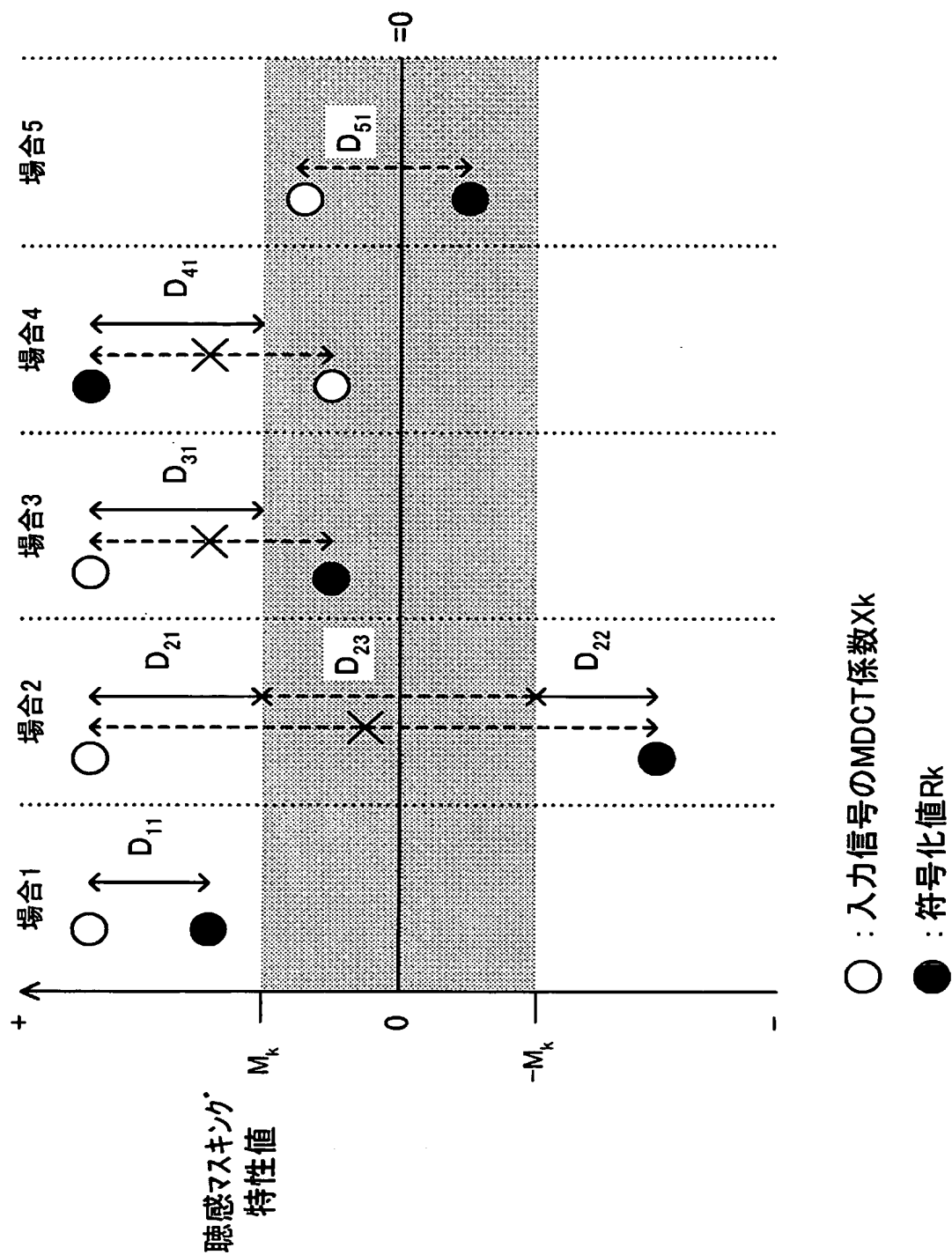
[図3]



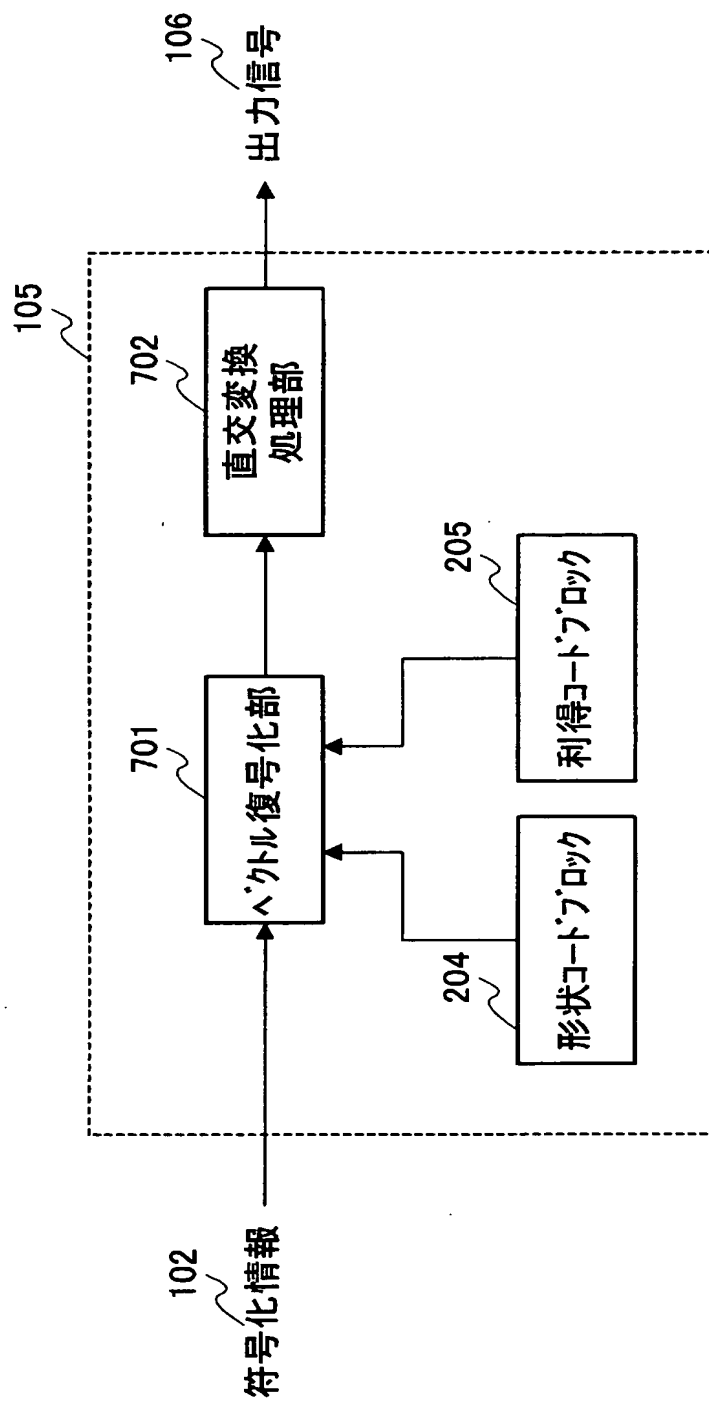
[図5]



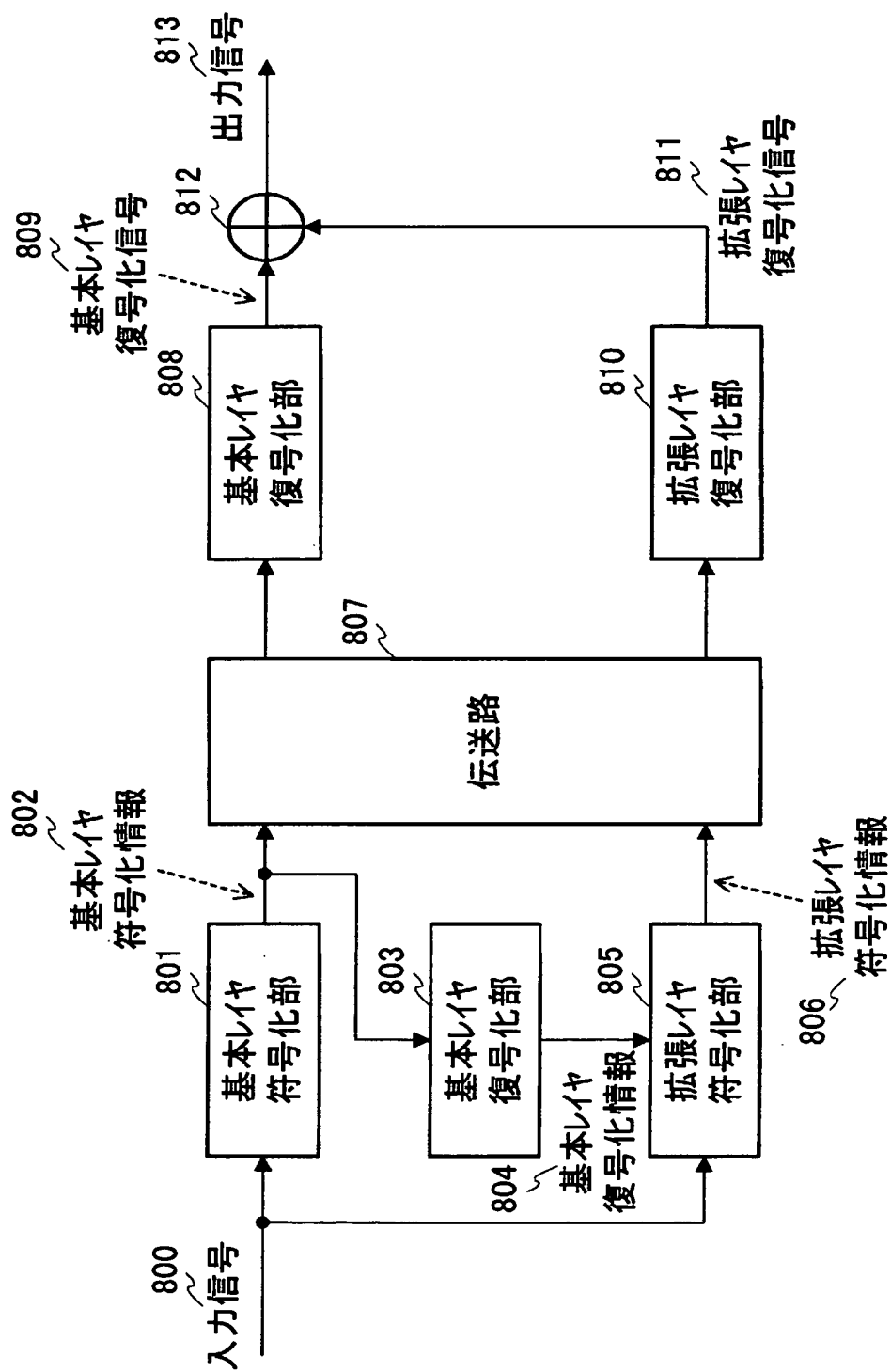
[図6]



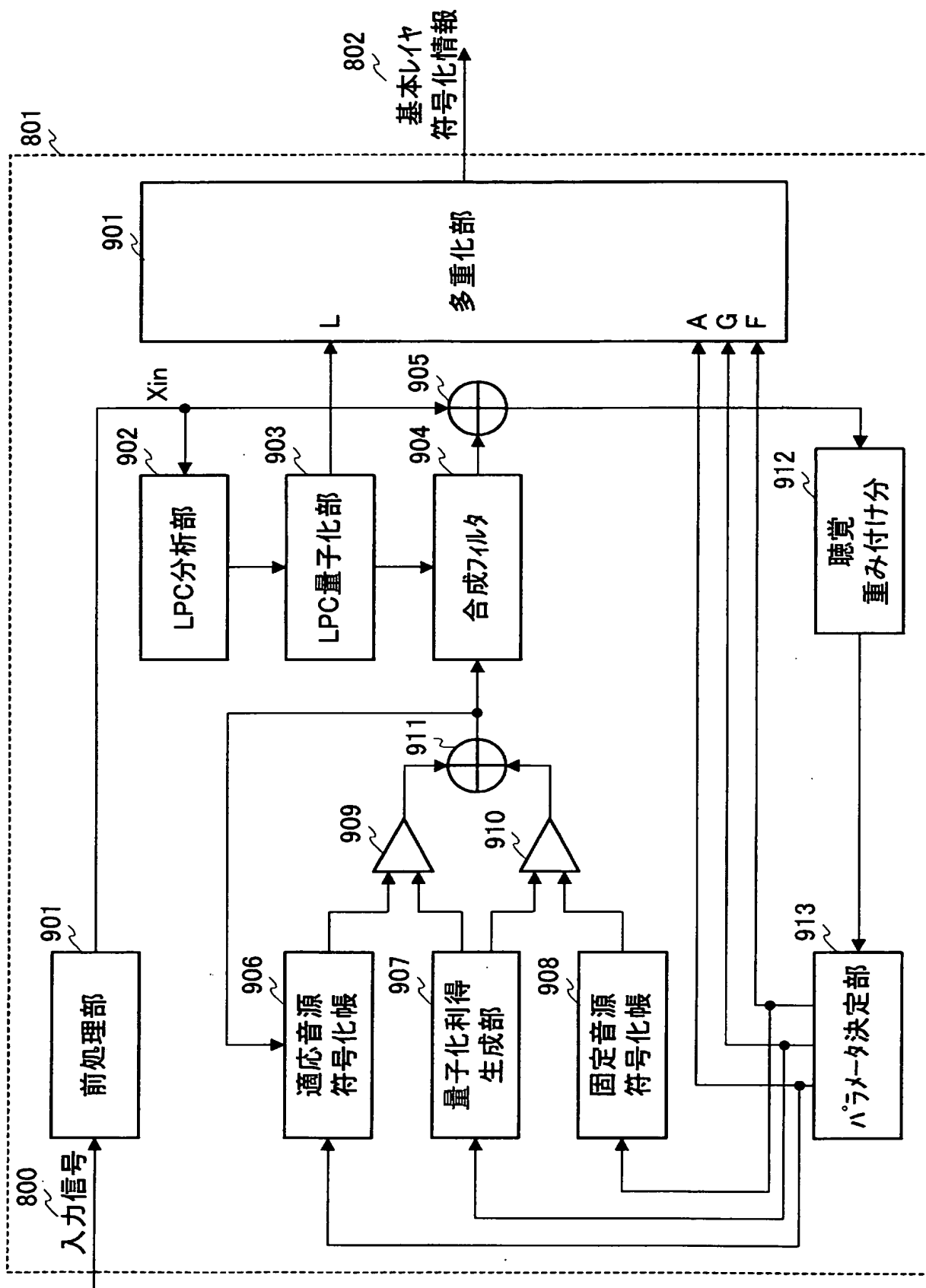
[図7]



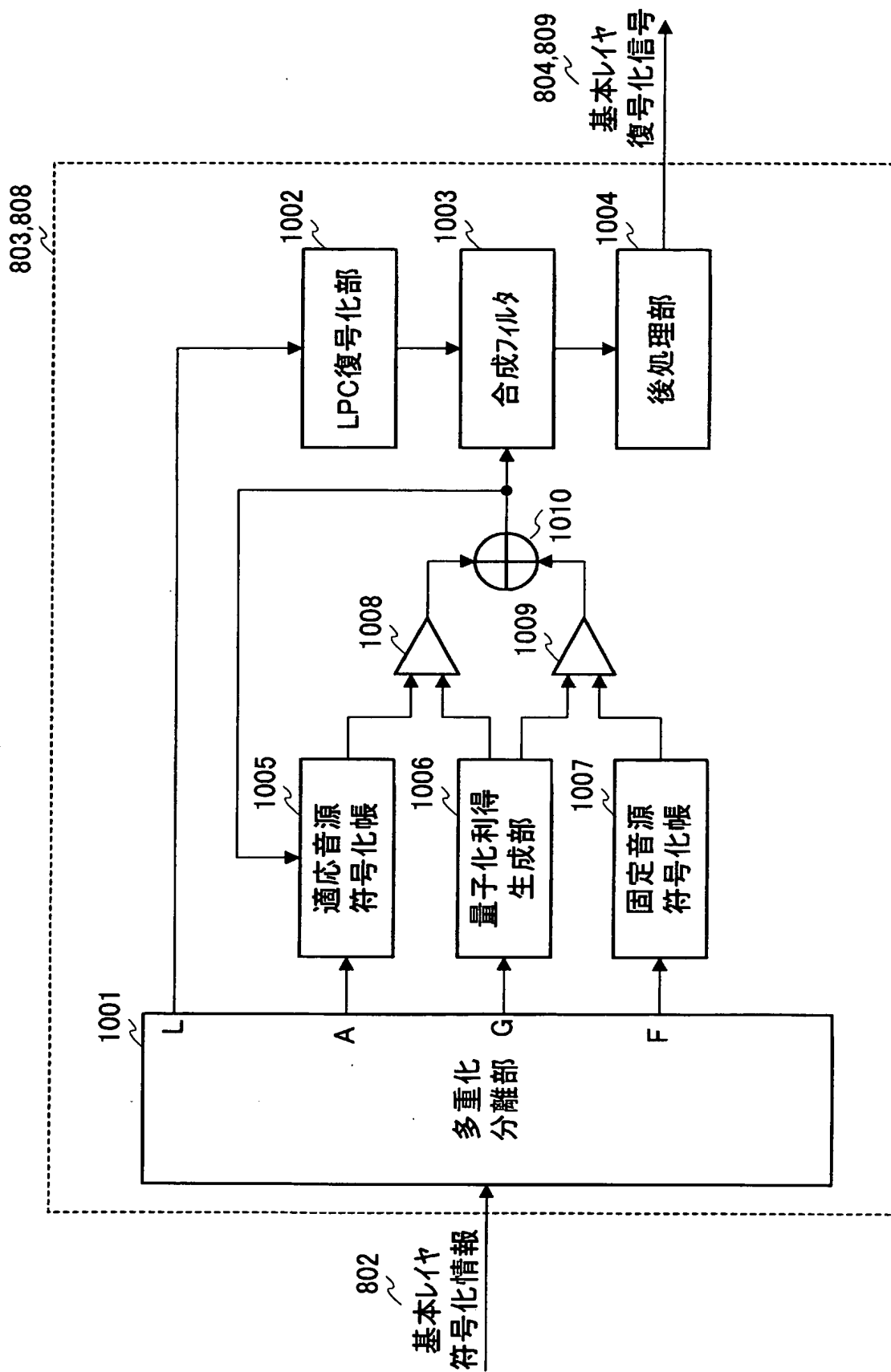
[図8]



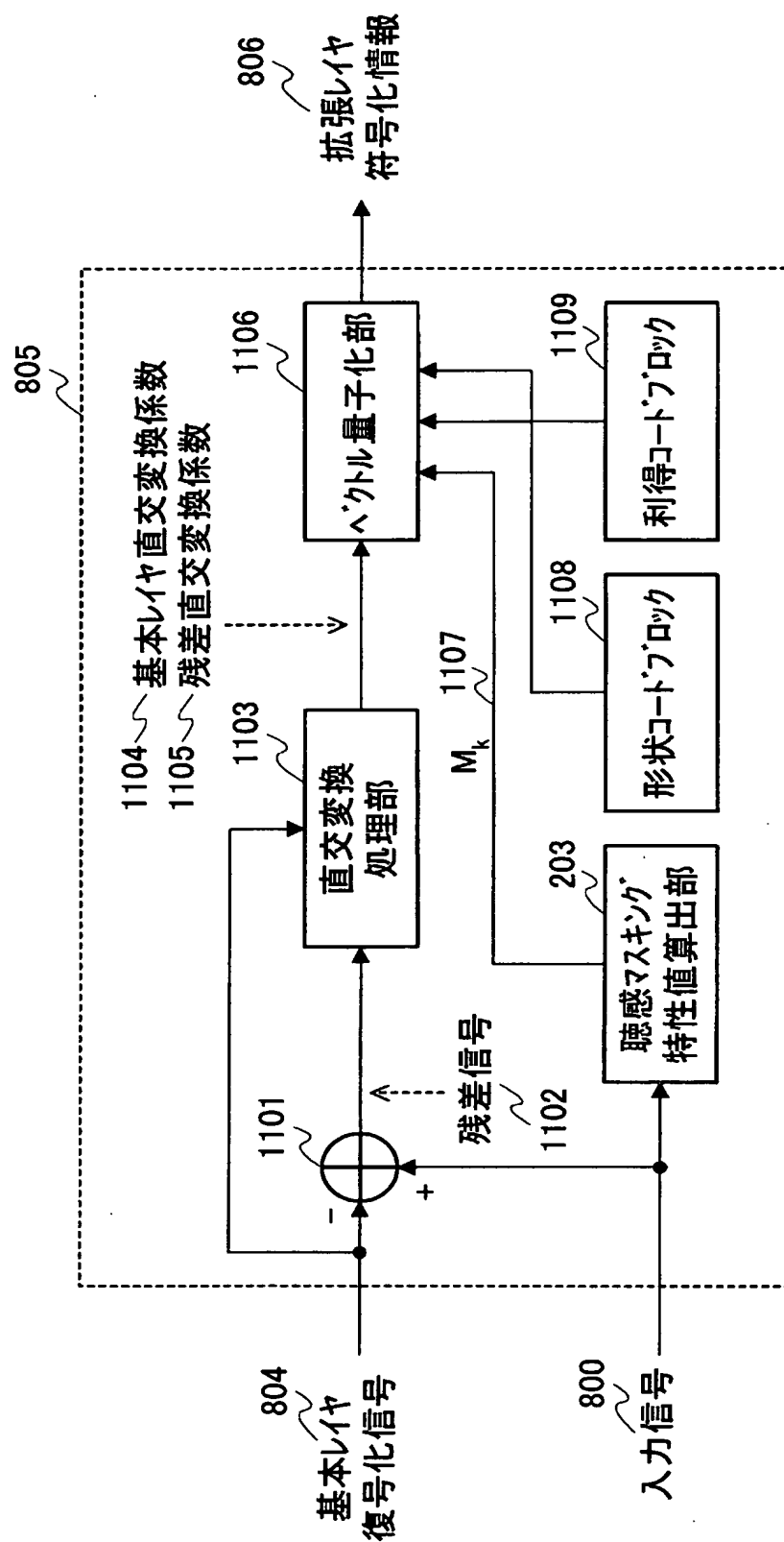
[図9]



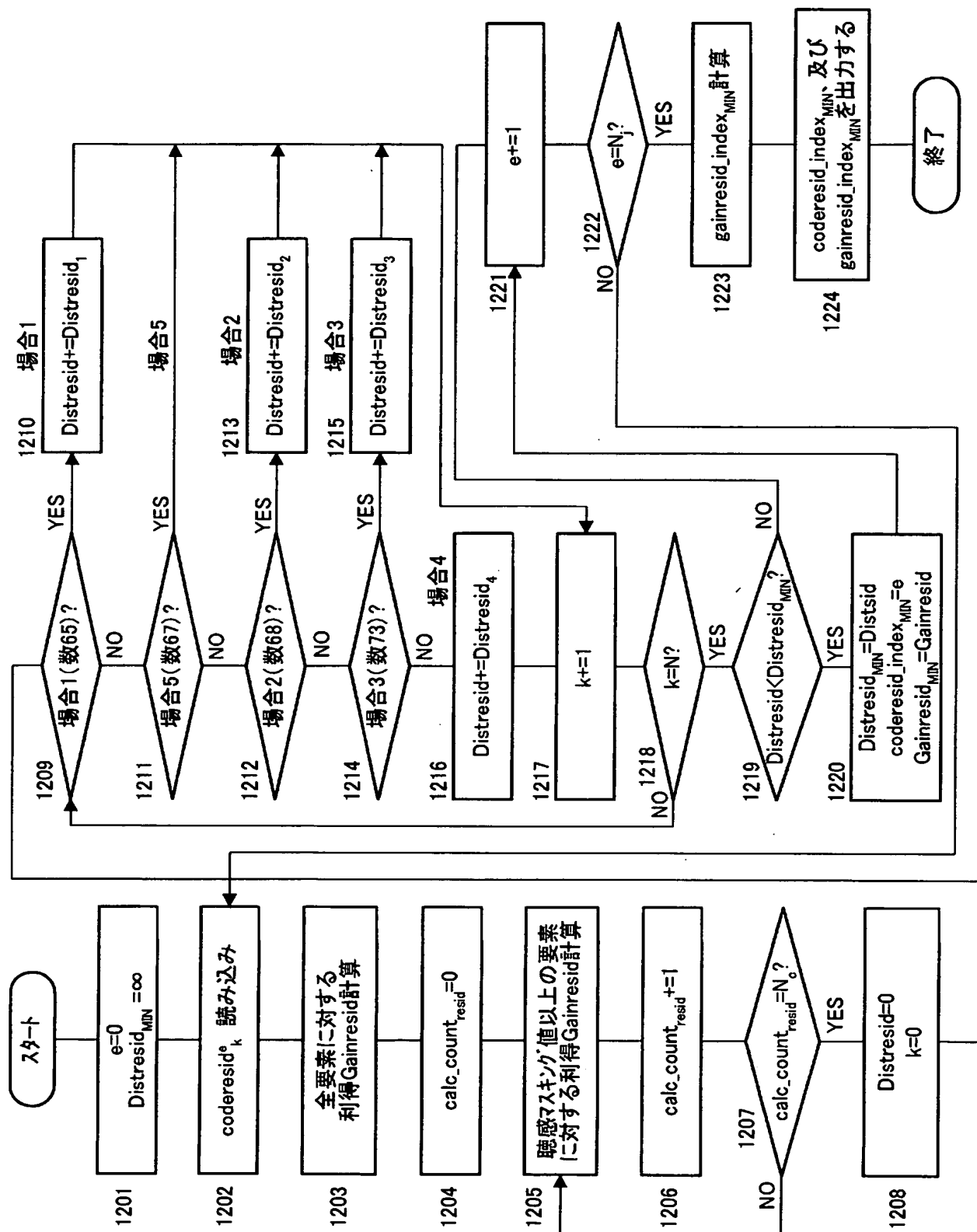
[図10]



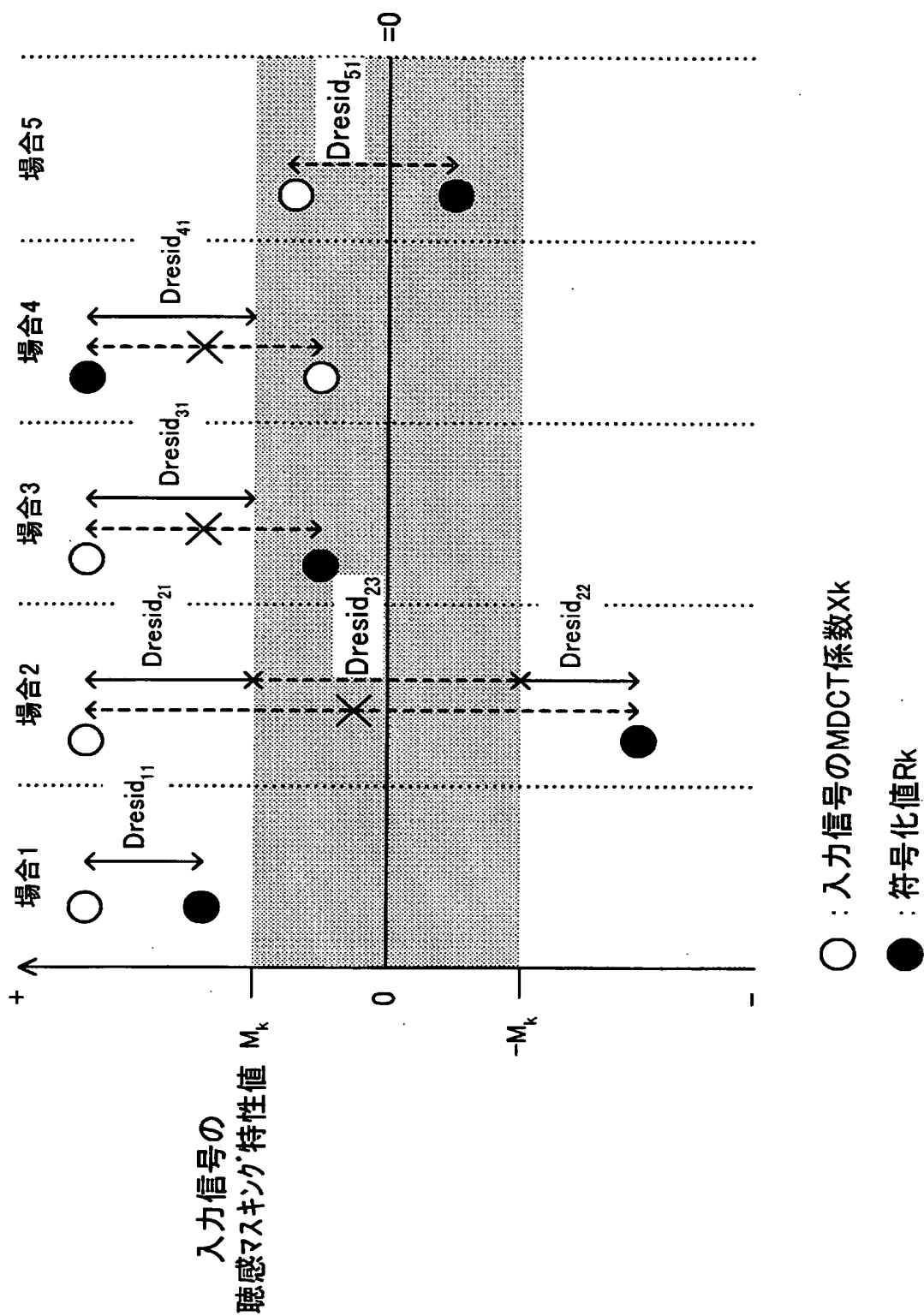
[図11]



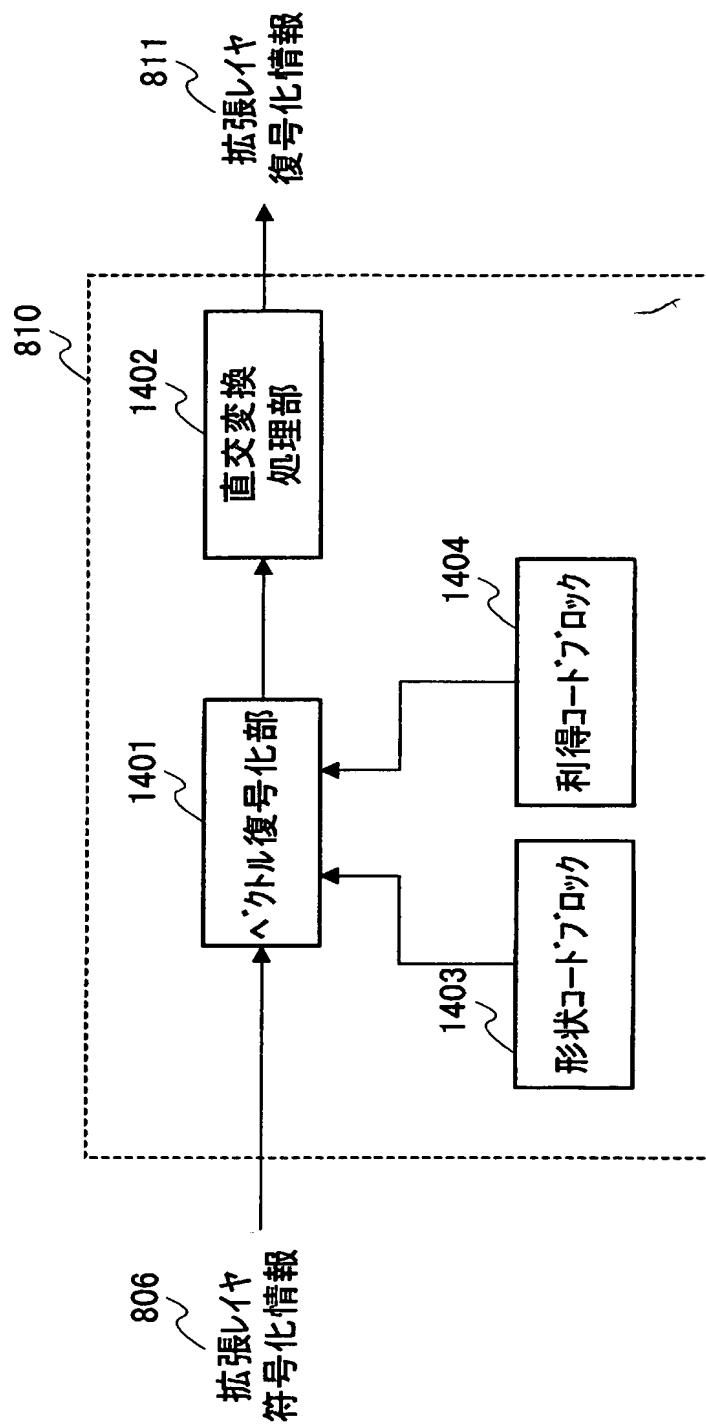
[図12]



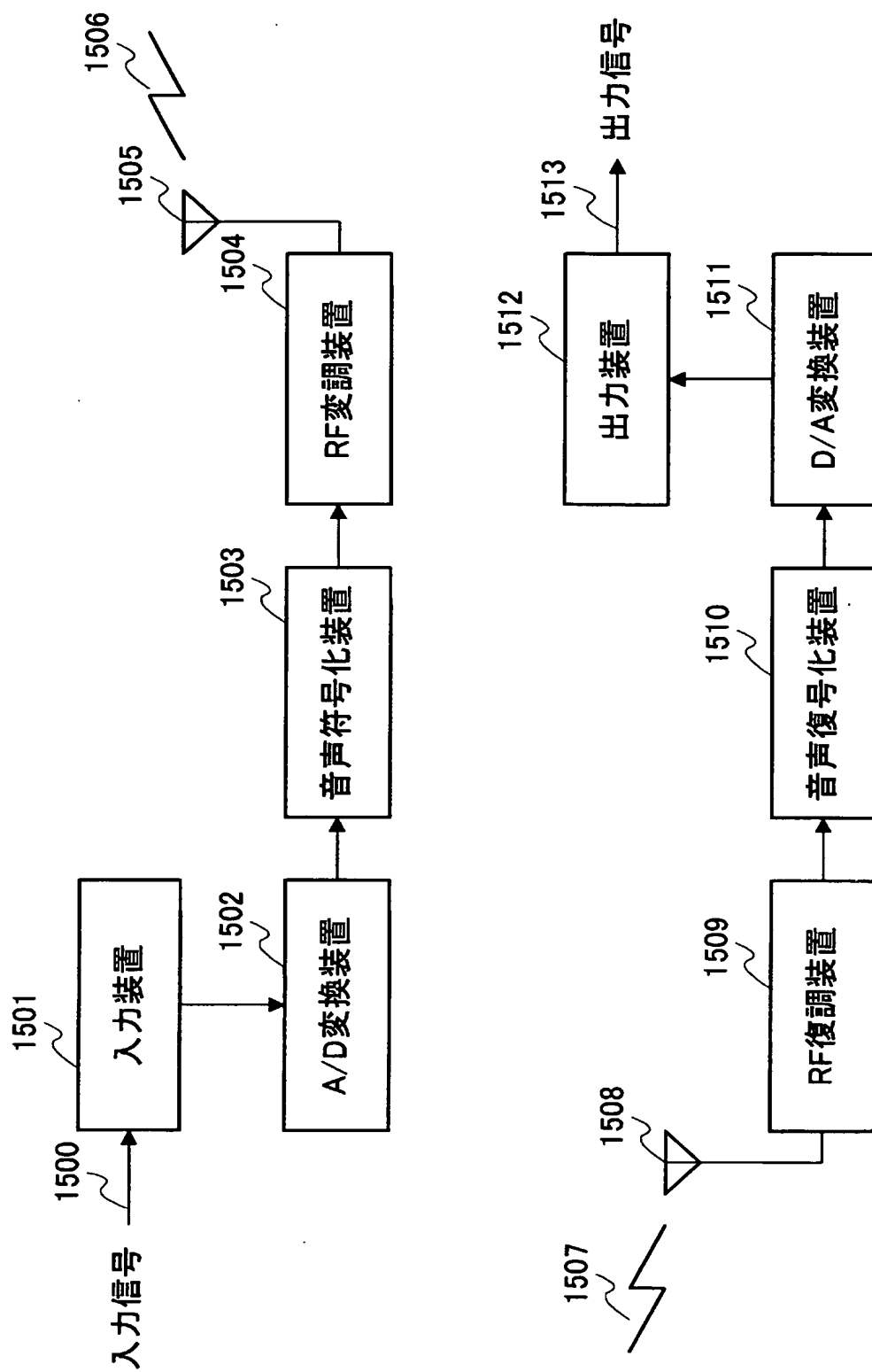
[図13]



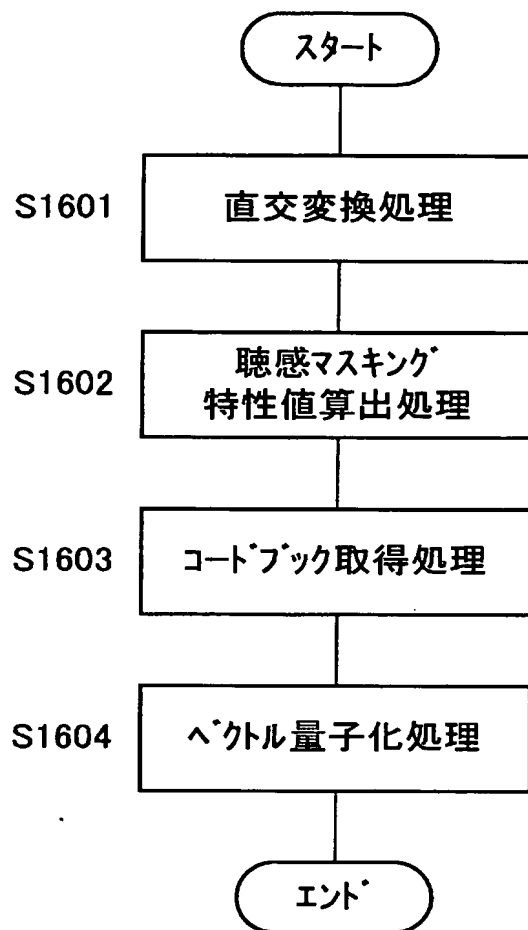
[図14]



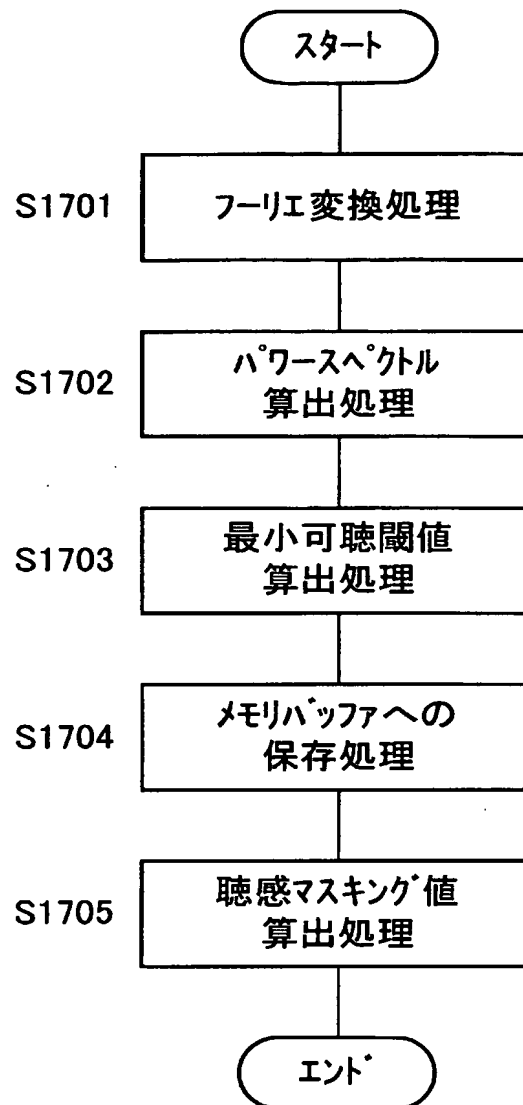
[図15]



[図16]



[図17]



INTERNATIONAL SEARCH REPORT

International application No.

PCT/JP2004/019014

A. CLASSIFICATION OF SUBJECT MATTER

Int.Cl⁷ G10L19/14

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

Int.Cl⁷ G10L19/00-19/14, H03M7/30-7/36

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Jitsuyo Shinan Koho	1922-1996	Jitsuyo Shinan Toroku Koho	1996-2005
Kokai Jitsuyo Shinan Koho	1971-2005	Toroku Jitsuyo Shinan Koho	1994-2005

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)

JICSTPlus FILE (JOIS)

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X Y	Tadashi YONEZAKI, Kiyohiro KANO, "Jikan Shuhasu Masking o Riyoshita Spectrum Horaku no Vector Ryoshika", The Acoustical Society of Japan (ASJ), Heisei 7 Nendo Shuki Kenkyu Happyokai Koen Ronbunshu -I-, 27 September, 1995 (27.09.95), pages 283 to 284	1, 4, 5, 8, 9 2, 6, 7
X Y	JP 8-123490 A (Matsushita Electric Industrial Co., Ltd.), 17 May, 1996 (17.05.96), Full text; all drawings (Family: none)	1, 4, 5, 8, 9 2, 6, 7

☒ Further documents are listed in the continuation of Box C.

☐ See patent family annex.

* Special categories of cited documents:	"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention
"A" document defining the general state of the art which is not considered to be of particular relevance	"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone
"E" earlier application or patent but published on or after the international filing date	"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art
"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)	"&" document member of the same patent family
"O" document referring to an oral disclosure, use, exhibition or other means	
"P" document published prior to the international filing date but later than the priority date claimed	

Date of the actual completion of the international search
11 March, 2005 (11.03.05)

Date of mailing of the international search report
29 March, 2005 (29.03.05)

Name and mailing address of the ISA/
Japanese Patent Office

Authorized officer

Facsimile No.

Telephone No.

INTERNATIONAL SEARCH REPORT

International application No.

PCT/JP2004/019014

C (Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
Y	JP 2003-323199 A (Matsushita Electric Industrial Co., Ltd.), 14 November, 2003 (14.11.03), Full text; all drawings & WO 2003/091989 A1	2, 6, 7
A	JP 7-160297 A (NEC Corp.), 23 June, 1995 (23.06.95), Full text; all drawings & US 005666465 A1 & EP 000658876 A3 & CA 002137757 A1	1-9
A	JP 2003-58196 A (Matsushita Electric Industrial Co., Ltd.), 08 February, 2003 (08.02.03), Par. Nos. [0068] to [0071]; all drawings & JP 11-330977 A & EP 000942411 A2 & CN 001240978 A	1-9
A	JP 11-327600 A (Matsushita Electric Industrial Co., Ltd.), 26 November, 1999 (26.11.99), Full text; all drawings & US 006311153 B1 & US 2001/0044727 A1 & EP 000907258 A1 & CA 002249792 A1	1-9

A. 発明の属する分野の分類 (国際特許分類 (IPC))

Int. Cl⁷ G10L19/14

B. 調査を行った分野

調査を行った最小限資料 (国際特許分類 (IPC))

Int. Cl⁷ G10L19/00-19/14
H03M 7/30- 7/36

最小限資料以外の資料で調査を行った分野に含まれるもの

日本国実用新案公報 1922-1996
 日本国公開実用新案公報 1971-2005
 日本国実用新案登録公報 1996-2005
 日本国登録実用新案公報 1994-2005

国際調査で使用した電子データベース (データベースの名称、調査に使用した用語)

JICSTPlusファイル (JOIS)

C. 関連すると認められる文献

引用文献の カテゴリー*	引用文献名 及び一部の箇所が関連するときは、その関連する箇所の表示	関連する 請求の範囲の番号
X Y	米崎正, 鹿野清宏, “時間周波数マスキングを利用したスペクトル包絡のベクトル 量子化”, 日本音響学会 平成7年度秋季研究発表会 講演論文集 -I-, 1995. 09. 27, p. 283-284	1, 4, 5, 8, 9 2, 6, 7
X Y	JP 8-123490 A (松下電器産業株式会社) 1996. 05. 17, 全文, 全図 (ファミリーなし)	1, 4, 5, 8, 9 2, 6, 7

☒ C欄の続きにも文献が列挙されている。☐ パテントファミリーに関する別紙を参照。

* 引用文献のカテゴリー

「A」 特に関連のある文献ではなく、一般的技術水準を示すもの
 「E」 国際出願日前の出願または特許であるが、国際出願日以後に公表されたもの
 「L」 優先権主張に疑義を提起する文献又は他の文献の発行日若しくは他の特別な理由を確立するために引用する文献 (理由を付す)
 「O」 口頭による開示、使用、展示等に言及する文献
 「P」 国際出願日前で、かつ優先権の主張の基礎となる出願

の日の後に公表された文献

「T」 国際出願日又は優先日後に公表された文献であって出願と矛盾するものではなく、発明の原理又は理論の理解のために引用するもの
 「X」 特に関連のある文献であって、当該文献のみで発明の新規性又は進歩性がないと考えられるもの
 「Y」 特に関連のある文献であって、当該文献と他の1以上の文献との、当業者にとって自明である組合せによって進歩性がないと考えられるもの
 「&」 同一パテントファミリー文献

国際調査を完了した日

11. 03. 2005

国際調査報告の発送日

29. 3. 2005

国際調査機関の名称及びあて先

日本国特許庁 (ISA/J P)
 郵便番号100-8915
 東京都千代田区霞が関三丁目4番3号

特許庁審査官 (権限のある職員)

山下 剛史

5 C

3 3 5 2

電話番号 03-3581-1101 内線 3541

C (続き) . 関連すると認められる文献		
引用文献の カテゴリー*	引用文献名 及び一部の箇所が関連するときは、その関連する箇所の表示	関連する 請求の範囲の番号
Y	JP 2003-323199 A (松下電器産業株式会社) 2003.11.14, 全文, 全図 & WO 2003/091989 A1	2, 6, 7
A	JP 7-160297 A (日本電気株式会社) 1995.06.23, 全文, 全図 & US 005666465 A1 & EP 000658876 A3 & CA 002137757 A1	1-9
A	JP 2003-58196 A (松下電器産業株式会社) 2003.02.08, 【0068】 - 【0071】, 全図 & JP 11-330977 A & EP 000942411 A2 & CN 001240978 A	1-9
A	JP 11-327600 A (松下電器産業株式会社) 1999.11.26 全文, 全図 & US 006311153 B1 & US 2001/0044727 A1 & EP 000907258 A1 & CA 002249792 A1	1-9